

WORLD INTELLECTUAL PROPERTY ORGANIZATION International Bureau



INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6:

C12N 15/49, C07K 14/16, A61K 39/21

A1

(43) International Patent Classification 6:

(11) International Patent Classification 6:

(12) International Patent Classification 6:

(13) International Patent Classification 6:

(14) International Patent Classification 6:

(15) International Patent Classification 6:

(15) International Patent Classification 6:

(16) International Patent Classification 6:

(17) International Patent Classification 6:

(18) International Patent Classification 6:

(19) International Patent Classification 6:

(19) International Patent Classification 6:

(10) International Patent Classification 6:

(11) International Patent Classification 6:

(11) International Patent Classification 6:

(12) International Patent Classification 6:

(13) International Patent Classification 6:

(14) International Patent Classification 6:

(15) International Patent Classification 6:

(16) International Patent Classification 6:

(17) International Patent Classification 6:

(18) International Patent Classification 6:

(18) International Patent Classification 6:

(19) International Patent Classification 6:

(19) International Patent Classification 6:

(10) International Patent Classification 6:

(10) International Patent Classification 6:

(10) International Patent Classification 6:

(11) International Patent Classification 6:

(12) International Patent Classification 6:

(13) International Patent Classification 6:

(14) International Patent Classification 6:

(15) International Patent Classification 6:

(16) International Patent Classification 6:

(17) International Patent Classification 6:

(18) Internation 6:

(18) Internation 6:

(18) Internation 6:

(18) Int

(11) International Publication Number:

WO 98/01564

(43) International Publication Date:

15 January 1998 (15.01.98)

(21) International Application Number:

PCT/US97/09690

(22) International Filing Date:

3 July 1997 (03.07.97)

(30) Priority Data:

08/676,737

8 July 1996 (08.07.96)

US

(60) Parent Application or Grant

(63) Related by Continuation

US Filed on 08/676,737 (CIP) 8 July 1996 (08.07.96)

(71) Applicant (for all designated States except US): GENENTECH, INC. [US/US]; 460 Point San Bruno Boulevard, South San Francisco, CA 94080 (US).

(72) Inventor; and

(75) Inventor/Applicant (for US only): BERMAN, Phillip, W. [US/US]; 95 Cheyene Point, Portola Valley, CA 94028 (US).

(74) Agents: HALIDAY, Emily, H. et al.; Skjerven, Morrill, MacPherson, Franklin & Friel, Suite 700, 25 Metro Drive, San Jose, CA 95110 (US). (81) Designated States: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GE, GH, HU, IL, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ARIPO patent (GH, KE, LS, MW, SD, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG).

Published

With international search report.

Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.

(54) Title: HIV ENVELOPE POLYPEPTIDES AND VACCINE

(57) Abstract

Oligonucleotide sequences encoding gp120 polypeptides from breakthrough isolates of vaccine trials using MN-rgp120 and the encoded gp120 polypeptides are provided. Use of the gp120 polypeptides from one or more of the isolates in a subunit vaccine, usually together with MN-rgp120, can provide protection against HIV strains that are sufficiently different from the vaccine strain (e.g.; MN-rgp120) that the vaccine does not confer protection against those strains. Antibodies induced by the polypeptides are also provided.

193

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albenia	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Annesia	7	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	Prance	LÜ	Luxemboure	SN	Senegal
ΑÜ	Australia	GA	Gabon	LV	Latvia	5 Z	Swaziland
AZ	Azerbeijan	GB	United Kingdom	MC	Mosaco	770	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav	TM	Turkmenistan
BF	Burkina Faso	GR	Greece		Republic of Macedonia	TR	Turkey
BG	Bulgaria	HU	Hungary	ML	Mali	11	Trinidad and Tobago
Ŋ	Benin	Œ	Ireland	MN	Mongolia	UA	Ukraine
BR	Brazil	1L	Israel	MR	Mauritania	UG	Uganda
BY	Belarus	E1	iceland	MW	Malawi	US	United States of America
CA	Canada	IT	Italy	MX	Mexico	UZ	Uzbekistan
CT	Central African Republic	JP	Japan	NE	Niger	VN	Viot Nam
CG	Congo	KE	Kenya	NL	Netherlands	YU	Yugoslavia
CH	Switzerland	KG	Kyngyzstan	NO	Norway	zw	Zimbebwe
a	Côte d'Ivoire	KP	Democratic People's	NZ	New Zealand		
CM	Cameroon		Republic of Korea	PL	Poland		
CN	China	KR	Republic of Korea	PT	Portugal		
CU	Cuba	KZ	Kazakstan	RO	Romania		
Œ	Caech Republic	ıc	Saint Lucia	RU	Russian Federation		
DE	Germany	u	Liechtenstein	SD	Sedan		
DK	Deamark	LK	Sri Lanks	SE	Sweden		
KE.	Estonia	LR	Liberia	SG	Singapore		

PCT/US97/09690 WO 98/01564

HIV ENVELOPE POLYPEPTIDES AND VACCINE

BACKGROUND OF THE INVENTION

Field of the Invention 5

10

15

This invention relates to HIV envelope polypeptides and vaccines containing the polypeptides.

Description of the Related Art

Acquired immunodeficiency syndrome (AIDS) is caused by a retrovirus identified as the human immunodeficiency virus (HIV). There have been intense efforts to develop a vaccine that induces a protective immune response based on induction of antibodies or cellular responses. Recent efforts have used subunit vaccines where an HIV protein, rather than attenuated or killed virus, is used as the immunogen in the vaccine for safety reasons. Subunit vaccines generally include gp120, the portion of the HIV envelope protein which is on the surface of the virus.

20 The HIV envelope protein has been extensively described, and the amino acid and nucleic acid sequences encoding HIV envelope from a number of HIV strains are known (Myers, G. et al., 1992. Human Retroviruses and AIDS. A compilation and analysis of 25 nucleic acid and amino acid sequences. Los Alamos National Laboratory, Los Alamos, New Mexico). The HIV envelope protein is a glycoprotein of about 160 kd (gp160) which is anchored in the membrane bilayer at its carboxyl terminal region. The N-terminal segment, 30 gp120, protrudes into the aqueous environment surrounding the virion and the C-terminal segment, gp41, spans the membrane. Via a host-cell mediated process, gp160 is cleaved to form gp120 and the integral membrane protein gp41. As there is no 35 covalent attachment between gp120 and gp41, free gp120

is sometimes released from the surface of virions and

infected cells.

10

15

20

25

30

35

The gp120 molecule consists of a polypeptide core of 60,000 daltons which is extensively modified by N-linked glycosylation to increase the apparent molecular weight of the molecule to 120,000 daltons. The amino acid sequence of gp120 contains five relatively conserved domains interspersed with five hypervariable domains. The positions of the 18 cysteine residues in the gp120 primary sequence, and the positions of 13 of the approximately 24 N-linked glycosylation sites in the gp120 sequence are common to all gp120 sequences. The hypervariable domains contain extensive amino acid substitutions, insertions and deletions. Sequence variations in these domains result in up to 30% overall sequence variability between gp120 molecules from the various viral isolates. Despite this variation, all gp120 sequences preserve the ability of the virus to bind to the viral receptor CD4 and to interact with gp41 to induce fusion of the viral and host cell membranes.

gp120 has been the object of intensive investigation as a vaccine candidate for subunit vaccines, as the viral protein which is most likely to be accessible to immune attack. At present, clinical trials using gp120 MN strain are underway. However, to date no human vaccine trial has been of sufficient size to confirm or refute vaccine efficacy.

The development of candidate HIV-1 vaccines is burdened by the lack of in vivo or in vitro models of HIV-1 infection that accurately approximate the conditions of natural infection in humans. Several candidate HIV-1 vaccines (Berman et al.; J. Virol. 7:4464-9 (1992); Haigwood et al.; J. Virol. 66:172-82 (1992); Salmon-Ceron et al.; AIDS Res. and Human Retroviruses 11:1479-86 (1995)} have been described that elicit broadly cross-reactive antibodies able to

neutralize a variety of diverse HIV-1 isolates in However, the relevance of in vitro assays to protective immunity in vivo is uncertain. Although several vaccines have provided chimpanzees with protection from challenge by homologous and heterologous strains of HIV-1, protection has not always correlated with in vitro neutralization assays carried out in T cell lines, or in lectin- and cytokine-activated peripheral blood mononuclear cells (PBMCs) [Berman et al.; Nature 345:622-5 (1990); Bruck et al.; Vaccine 12(12):1141-8 (1994); El-Amad et al.; AIDS 9:1313-22 (1995); Girard et al.; J. Virol. 69:6239-48 (1995); and Fulz et al; Science 256:1687-1690 (1992)]. While successful protection of chimpanzees is encouraging and has historically proved to be a reliable indicator of vaccine efficacy, the conditions of infection in all experimental models of HIV-1 infection differ significantly from natural infection in humans.

10

15

Experimental HIV-1 infection in vivo and in vitro 20 both suffer from the limitation that the in vitro amplification of HIV-1, which is required to prepare virus stocks for in vitro or in vivo infectivity experiments, imposes a genetic selection that results in a spectrum of virus quasi-species that differ from 25 the spectrum of variants present in the clinical specimens used to establish the culture [Kusumi et al.; J. Virol. 66:875 (1992); Meyerhans et al.; Cell 58:901-10 (1989)]. Because of these uncertainties, and even greater uncertainties related to the amount of 30 virus transmitted, the site and cell type involved in initial replication, and the kinetics of virus dissemination, the ability of currently available in vitro or in vivo assays to reliably predict vaccine efficacy is questionable. 35

One of the candidate HIV-1 vaccines that have

entered human clinical trials is recombinant gp120 prepared in Chinese hamster ovary (CHO) cells from the MN strain of HIV-1 (MN-rgp120) (Berman et al.; J. Virol. 7:4464-9 (1992)). To date, approximately 499 adults have participated in Phase 1 and 2 immunogenicity and safety trials of this vaccine. data collected thus far suggest that MN-rgp120 is safe, immunogenic, and elicits high titers of neutralizing antibodies in greater than 95% of individuals immunized according to a 0, 1, and 6 month immunization schedule 10 [Belshe et al.; JAMA 272(6):475-80 (1994); McElrath; Seminars in Cancer Biol. 6:1-11 (1995)]. However, during the course of these trials, nine vaccinees who received MN-rgp120 have become infected with HIV-1 15 through high risk behavior. Small trials, such as these, in populations with low rates of infection and minimally sized placebo control groups do not have sufficient statistical power to confirm or refute vaccine efficacy.

However, effective vaccines based on gp120 or another HIV protein for protection against additional strains of HIV are still being sought to prevent the spread of this disease.

25 Description of the Background Art

30

Recombinant subunit vaccines are described in Berman et al., PCT/US91/02250 (published as number WO91/15238 on 17 October 1991). See also, e.g.
Hu et al., Nature 328:721-724 (1987) (vaccinia virus-HIV envelope recombinant vaccine); Arthur et al.,
J. Virol. 63(12): 5046-5053 (1989) (purified gp120); and Berman et al., Proc. Natl. Acad. Sci. USA 85:5200-5204 (1988) (recombinant envelope glycoprotein gp120).

Numerous sequences for gp120 are known. The sequence of gp120 from the IIIB substrain of HIV-1

referred to herein is that determined by Muesing et al., "Nucleic acid structure and expression of the human AIDS/lymphadenopathy retrovirus, Nature 313:450-458 (1985). The sequences of gp120 from the NY-5, Jrcsf, Z6, Z321, and HXB2 strains of HIV-1 are listed by Myers et al., "Human Retroviruses and AIDS; A compilation and analysis of nucleic acid and amino acid sequences," Los Alamos National Laboratory, Los Alamos, New Mexico (1992). The sequence of the Thai isolate A244 is provided by McCutchan et al., "Genetic Variants of HIV-1 in Thailand," AIDS Res. and Human Retroviruses 8:1887-1895 (1992). The MN_{10M} clone is described by Gurgo et al., "Envelope sequences of two new United States HIV-1 isolates, " Virol. 164: 531-536 (1988). As used herein, MN, MN-rgp120, the MN clone or isolate The MN_{GNE} amino acid sequence is refers to MNGNE. Sequence ID No. 29.

Each of the above-described references is incorporated herein by reference in its entirety.

20

25

30

35

15

5

10

Summary of the Invention

Oligonucleotide sequences encoding gp120 polypeptides from breakthrough isolates of vaccine trials using MN-rgp120 and the encoded gp120 polypeptides are provided. Use of the gp120 polypeptides from one or more of the isolates in a subunit vaccine, usually together with MN-rgp120, can provide protection against HIV strains that are sufficiently different from the vaccine strain (e.g.; MN-rgp120) that the vaccine does not confer protection against those strains. Antibodies induced by the polypeptides are also provided.

Brief Description of the Drawings

Figure 1 illustrates the kinetics of antibody response to MN-rgp120 in vaccinees infected with HIV-1.

Sera were collected at the time points indicated and assayed for antibodies reactive with MN-rgp120 (open circles) or a synthetic peptide derived from the V3 domain of MN-rgp120 (closed circles). Arrows indicate dates of injection. Plus sign indicates the first time HIV-1 infection was detected. Shaded area indicates data collected after HIV-1 infection. Data from vaccinee C6 is shown in panel A; C8 in panel B; C7, panel C; C11, panel D; C10, panel E; C17, panel F; and C15, panel G.

10

15

20

25

30

35

Figure 2 illustrates the kinetics of CD4 blocking antibody response in vaccinees infected with HIV-1. Sera were collected at the time points indicated and assayed for antibodies able to block the binding of [125I]-labeled MN-rgp120 to cell surface CD4. Arrows indicate dates of injection. Plus sign indicates the first time HIV-1 infection was detected. Shaded area indicates data collected after HIV-1 infection. Data from vaccinee C6 is shown in panel A; C8 in panel B; C7, panel C; C11, panel D; C10, panel E; C17, panel F; and C15, panel G.

Figure 3 illustrated predicted amino acid sequences of envelope glycoproteins (gp120) from breakthrough viruses. Proviral DNA sequences were amplified by PCR from PBMCs and cloned into the PRK5 expression plasmid. Two clones from each infected vaccinee were sequenced from double stranded plasmid DNA. Sequence numbering is with reference to the initiator methionine residue of gp120. For the purpose of comparison, the sequences shown begin at amino acid 12 of the mature, fully processed, envelope glycoproteins (corresponding to position 41 of the gp120 open reading frame). Shaded areas indicate sequences at neutralizing epitopes, dark boxes indicate polymorphisms thought to be important for the binding of virus neutralizing MAbs reactive with MN-rgp120.

Conserved (C) regions and variable (V) regions are indicated above the sequences. Boxes indicate sequence homologies and polymorphisms.

5

10

15

20

25

30

35

Figure 4 illustrates immunoprecipitation of recombinant gp120 prepared from breakthrough viruses. Recombinant gp120s from the seven breakthrough viruses were prepared by transient transfection of 293s cells. Cells were metabolically labeled with 35S methionine and growth conditioned cell culture supernatants were immunoprecipitated with polyclonal antisera to MN-rgp120. Immunoprecipitates were resolved by SDS-PAGE and visualized by autoradiography. C8 lanes a and b correspond to clones C8.3 and C8.6; C6 lanes a and b correspond to clones C6.1 and C6.5; C7 lanes a and b correspond to clones C7.2 and C7.10; C17 lanes a and b correspond to C17.1 and C17.3; C11 lanes a and b correspond to clones C11.5 and C11.7; C10 lanes a and b correspond to clones C10.5 and C10.7; C15 lanes a and b correspond to clones C15.2 and C15.3.

Figure 5 illustrates binding of monoclonal antibodies to recombinant gp120 from breakthrough viruses. Growth-conditioned cell culture supernatants were collected from 293s cells transiently transfected with plasmids directing the expression of breakthrough virus envelope glycoproteins. The relative rgp120 concentrations were determined by ELISA using MAb 5B6 specific for the HSV-1 glycoprotein D flag epitope at the amino terminus of all of the rgp120 variants described herein. The resulting rgp120 preparations were captured onto wells of microtiter plates coated with a polyclonal antibody specific for a conserved sequence in the C-terminus of gp120. The binding of virus neutralizing monoclonal antibodies reactive with gp120 was determined by ELISA. A, binding by MAb (5B6) specific for the HSV-1 glycoprotein D flag epitope; B, binding by MAb (1034) against the V3 domain of

MN-rgpl20; C binding by MAb (50.1) raised against a synthetic peptide corresponding to the V3 domain of MN-rgpl20; D, binding by a human MAb (15e) known to block the binding of gpl20 to CD4.

5

10

Figure 6 depicts the mature envelope glycoprotein (gp120) from the MN clone of the MN strain of HIV-1 (SEQ. ID. NO. 29). Hypervariable domains are indicated in bold, and the V and C regions are indicated (according to Modrow et al., J. Virology 61(2):570 (1987). Potential glycosylation sites are marked with a (*).

Detailed Description of the Invention

The present invention provides gp120 polypeptides

from breakthrough isolates of HIV vaccine trials.

Novel oligonucleotide sequences encoding gp120 from

breakthrough isolates which can be used to express

gp120 are also provided. Use of gp120 polypeptides

from one or more of the isolates in a subunit vaccine,

usually together with MN-rgp120, can provide protection

against HIV strains that are sufficiently different

from the vaccine strain (e.g.; MN-rgp120) that the

vaccine does not confer protection against those

strains.

In one embodiment, the vaccine is based on the use of the MN-rgp120 polypeptide (Sequence ID No. 29) and gp120 polypeptides from MN-like viruses that include neutralizing epitopes that are not present in the initial vaccine strain, and are sufficiently different from those of the vaccine strain, to have been able to cause HIV-1 infections in MN-rgp120 vaccinated individuals (i.e.; to result in breakthrough infections). Use of the initial vaccine strain empirically determines the viruses present in the population that contain additional neutralizing epitopes sufficiently different from those of the

vaccine strain to escape protection induced by the vaccine strain. Use of an initial representative gp120 polypeptide in a vaccine acts as a sieve so that viruses that are not effectively protected against by the vaccine strain breakthrough the vaccine, empirically resulting in determination of additional strains in a given geographic region that are not protected against by the initial vaccine strain. Use of gp120 from those breakthrough isolates complements the vaccine isolate by providing additional neutralizing epitopes not present in the initial vaccine strain, therefore creating a more complete vaccine that confers protection against multiple different virus strains in the region.

5

10

15

20

25

30

35

Prior HIV-1 vaccine strategies were based on selection of appropriate candidate vaccine polypeptides based on homology alignment studies. However, since some of the neutralizing epitopes are conformationdependent and the location of all of these epitopes is not known, this approach necessarily cannot determine all of the neutralizing epitopes that should be included in a vaccine for a particular region. In contrast, the present approach uses a selected representative strain and empirically determines strains that are sufficiently different and therefore breakthrough the barrier of protection provided by the initial vaccination program. Those strains can be included in the vaccine to confer more complete protection from HIV strains in the region. addition, those strains can be used alone to confer protection against the breakthrough virus.

In another embodiment, the invention comprises a vaccine containing a first HIV gp120 polypeptide sequence and a breakthrough isolate HIV gp120 polypeptide sequence from a vaccinee vaccinated with a vaccine including the first HIV gp120 polypeptide

sequence, the HIV gp120 polypeptide sequences being in a suitable carrier. Fragments of one or both HIV gp120 polypeptide sequences can be substituted for one or both of the corresponding HIV gp120 polypeptide sequences.

5

10

15

20

25

30

35

Preferably, the first gp120 polypeptide sequence contains neutralizing epitopes found in one or more gp120 polypeptides present in isolates from the geographical region where the initial vaccine (i.e., the vaccine that gives rise to the breakthrough isolate) is administered. More preferably, the first gp120 polypeptide sequence contains at least one of the more common neutralizing epitopes for the region, and most preferably the first gp120 polypeptide sequence contains at least one of the three most common neutralizing epitopes.

gp120 polypeptide sequences suitable for use as the first gp120 polypeptide sequence include gp120 MN, the Thai isolate A244 sequence (hereinafter "gp120 A244"), gp120 MN-GNE6 (Sequence ID No. 31; also known in the art as "gp120 GNE6"), and gp120 MN-GNE8 (Sequence ID No. 33; also known in the art as "gp120 GNE8"), and the like. gp120 MN, gp120 MN-GNE6, and gp120 MN-GNE8 are especially preferred for use as the first gp120 polypeptide sequence in initial vaccines for North America. gp120 A244 is especially preferred for use as the first gp120 polypeptide sequence in initial vaccines for North America. gp120 polypeptide sequence in initial vaccines for Thailand.

In a variation of this embodiment, the vaccine includes two different (i.e., first and second) gp120 polypeptide sequences, or fragments thereof, in combination with a breakthrough isolate HIV gp120 polypeptide sequence. The latter can be from a vaccinee vaccinated with either or both of the first and second HIV gp120 polypeptide sequences.

Exemplary vaccines include those containing

combinations of gp120 MN, gp120 A244, gp120 MN-GNE6 (Sequence ID No. 31), and gp120 MN-GNE8 (Sequence ID No. 33). Combinations of gp120 MN and gp120 A244 or gp120 MN-GNE8 (Sequence ID No. 33) with a breakthrough isolate HIV gp120 polypeptide sequence are especially preferred.

In vaccines containing gp120 MN, the breakthrough isolate HIV gp120 polypeptide sequence can be an HIV gp120 polypeptide sequence selected from the group consisting of Sequence ID Nos. 2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, and 28, and fragments thereof.

10

15

20

25

30

35

The term "subunit vaccine" is used herein, as in the art, to refer to a viral vaccine that does not contain virus, but rather contains one or more viral proteins or fragments of viral proteins. As used herein, the term "multivalent", means that the vaccine contains gp120 from at least two HIV isolates having different amino acid sequences.

The term "breakthrough isolate" or "breakthrough virus" is used herein, as in the art, to refer to a virus isolated from a vaccinee.

The terms "amino acid sequence", "polypeptide sequence", and "polypeptide" are used interchangeably herein as in the art, as are the terms "nucleic acid sequence", "nucleotide sequence", and "oligonucleotide".

Polypeptides from Breakthrough Isolates

The gp120 polypeptides of this invention correspond to the amino acid sequences of seven breakthrough isolates which are illustrated below in Table 1. A polypeptide of this invention includes an HIV gp120 amino acid sequence illustrated in Table 1 (Sequence ID Nos. 1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, and 27) and fragments thereof. The polypeptides of this invention can include fused

sequences from two or more HIV gp120 or gp160 amino acid sequences.

5

10

15

20

25

30

35

The polypeptide can also be joined to another viral protein, such as a flag epitope amino acid sequence. The term "flag epitope" is used herein, as in the art, to denote an amino acid sequence that includes an epitope recognized by a monoclonal antibody. Flag epitopes facilitate using single monoclonal antibody affinity purification of a plurality of different recombinant proteins, each having the flag epitope recognized by the monoclonal antibody. Numerous amino acid sequences can function as flag epitopes. The N-terminal sequences of Herpes Simplex Virus Type 1 (HSV-1) glycoprotein D (gD-1) is conveniently used as the flag epitope and its use is described in detail in the examples. The flag epitope is conveniently fused to the N terminus of the HIV gp120 polypeptide sequence. Alternatively, however, monoclonal antibodies that recognize neutralizing epitopes in the rgp120 sequences can be used to affinity purify the amino acid sequences, and a flag epitope can be omitted.

In addition, various signal sequences can be joined to a polypeptide of this invention. Although rgp120 is secreted to some extent in HIV cultures, the amount of the envelope glycoprotein released from (secreted by) the host cells varies widely from strain to strain. Various signal sequences can be introduced into the polypeptide by joining a nucleotide sequence encoding the signal sequence to the nucleotide sequence encoding the rgp120 to facilitate secretion of rgp120 from the cells. For example, Chiron HIV gp120 polypeptides include a signal sequence from tissue plasminogen activator (TPA) that provides good secretion of rgp120. Additional signal sequences are well known and include the N-terminal domain of murine

leukemia virus surface protein gp70 described by Kayman et al., J. Virol. 68:400-410 (1984).

Table 1 illustrates the nucleotide and deduced amino acid sequences for two clones of each the seven breakthrough isolates of this invention. The clones are: C6.1; C6.5; C8.3; C8.6; C15.2; C15.3; C7.2; C7.10; C11.5; C11.7; C10.5; C10.7; C17.1; and C17.3. These sequences are SEQ. ID. NOs. 1-28, the first sequence number for each clone being the nucleotide sequence and the second being the amino acid sequence. The amino acid sequence for MN and the nucleotide and deduced amino acid sequences for MN-GNE6 and MN-GNE8 are illustrated in the sequence listing hereinafter. In the listing for MN-GNE6, a stop codon appears at amino acid residue position 51. This stop codon can be replaced with a codon encoding the corresponding amino acid from MN or MN-GNE8 or another isolate.

10

15

TABLE 1

CLONE C6.1 GGG GTA CCT GTG TGG AAG GAA GCA ACC ACC ACT CTA 36 Gly Val Pro Val Trp Lys Glu Ala Thr Thr Leu 5 10 TTT TGT GCA TCA GAT GCT AAA GCA TAT GAC ACA GAG GTG 75 Phe Cys Ala Ser Asp Ala Lys Ala Tyr Asp Thr Glu Val 15 20 CAT AAT GTT TGG GCC ACA CAT GCT TGT GTA CCC ACA GAC 114 His Asn Val Trp Ala Thr His Ala Cys Val Pro Thr Asp 10 30 35 CCA AAC CCA CAA GAA ATG GTA TTG GAA AAT GTG ACA GAA 153 Pro Asn Pro Gln Glu Met Val Leu Glu Asn Val Thr Glu 15 40 45 GAT TTT AAC ATG TGG AAA AAT GAC ATG GTA GAA CAG ATG 192 Asp Phe Asn Met Trp Lys Asn Asp Met Val Glu Gln Met 55 60 CAT GAG GAT ATA ATC AGT TTA TGG GAT CAA AGC CTA AAA 231 His Glu Asp Ile Ile Ser Leu Trp Asp Gln Ser Leu Lys 20 70 CCA TGT GTA AAA TTA ACC CCA CTC TGT ATT ACT TTA AAT 270 Pro Cys Val Lys Leu Thr Pro Leu Cys Ile Thr Leu Asn 85 80 25 TGC ACC AAT TGG AAG AAG AAT GAT ACT AAA ACT AAT AGT 309 Cys Thr Asn Trp Lys Lys Asn Asp Thr Lys Thr Asn Ser 95 AGT AGT ACT ACA ACT AAT AAT AGT AGT GCT ACA GCT AAT 348 Ser Ser Thr Thr Asn Asn Ser Ser Ala Thr Ala Asn 30 110 AGT AGT AGT ACT ACA ACT AAT AGT AGT TGG GGA GAG ATA 387 Ser Ser Ser Thr Thr Thr Asn Ser Ser Trp Gly Glu Ile 125 120 AAG GAG GGA GAA ATA AAG AAC TGC TCT TTC AAT ATC ACC 426 Lys Glu Gly Glu Ile Lys Asn Cys Ser Phe Asn Ile Thr 35 135 ACA AGC ATA AGA GAC AAG GTG AAG AAA GAA TAT GCA CTT 465 Thr Ser Ile Arg Asp Lys Val Lys Lys Glu Tyr Ala Leu 150 145 40 TTT TAT AGC CTT GAT GTA GTA CCA ATA GAA AAT GAT AAT 504 Phe Tyr Ser Leu Asp Val Val Pro Ile Glu Asn Asp Asn 165 160 ACT AGC TAT AGG TTG AGA AGT TGT AAC ACC TCA GTC ATT 543 Thr Ser Tyr Arg Leu Arg Ser Cys Asn Thr Ser Val Ile 45 175 170 ACA CAA GCC TGT CCA AAG GTA ACT TTT GAG CCA ATT CCC 582 Thr Gln Ala Cys Pro Lys Val Thr Phe Glu Pro Ile Pro 190 185 ATA CAT TAT TGT ACC CCG GCT GGT TTT GCG ATT CTG AAG 621 Ile His Tyr Cys Thr Pro Ala Gly Phe Ala Ile Leu Lys 50 200 TGT AGA GAT AAA AAG TTC AAT GGA ACA GGA CCA TGC AAA 660 Cys Arg Asp Lys Lys Phe Asn Gly Thr Gly Pro Cys Lys 215 210 AAT GTT AGC ACA GTA CAA TGT GCA CAT GGA ATT AAG CCA 699 Asn Val Ser Thr Val Gln Cys Ala His Gly Ile Lys Pro 225 230 GTA GTG TCA ACT CAA CTG CTG TTA AAT GGC AGC CTA GCA 738 Val Val Ser Thr Gln Leu Leu Leu Asn Gly Ser Leu Ala 60 235 240 GAA GAA GAG GTA ATA ATT AGA TCT GCC AAT TTC TCA AAC 777 Glu Glu Glu Val Ile Ile Arg Ser Ala Asn Phe Ser Asn

```
AAT GCT AAA ATC ATA ATA GTA CAG TTG AGG GAA CCT GTA 816
      Asn Ala Lys Ile Ile Ile Val Gln Leu Arg Glu Pro Val
                         265
                                              270
      GAA ATT AAT TGT ACA AGA CCC AGC AAC AAT ACA ATA AAA 855
      Glu Ile Asn Cys Thr Arg Pro Ser Asn Asn Thr Ile Lys
                                  280
     GGT ATA CAC ATA GGA CCA GGG AGA GCA TTT TAT GCA ACA 894
     Gly Ile His Ile Gly Pro Gly Arg Ala Phe Tyr Ala Thr
                                         295
                    290
     GGA GAC ATA CGA GGA GAT ATA AGA CAA GCA CAT TGT AAC 933
10
     Gly Asp Ile Arg Gly Asp Ile Arg Gln Ala His Cys Asn
                             305
         300
      ATT AGT GGA GCA AAA TGG AAT AAC ACT TTA AAG AAG GTA 972
      Ile Ser Gly Ala Lys Trp Asn Asn Thr Leu Lys Lys Val
                                      320
                  315
15
     GTT AAA AAA TTA AAA GAA CAA TTT CCA AAT AAA ACA ATA 1011
      Val Lys Lys Leu Lys Glu Gln Phe Pro Asn Lys Thr Ile
                          330
                                               335
      325
      GTC TTT AAC CAT TCC TCA GGA GGG GAC CCA GAA ATT GTA 1050
      Val Phe Asn His Ser Ser Gly Gly Asp Pro Glu Ile Val
20
                                  345
      ATG CAC AGT TTT AAT TGT CAA GGG GAA TTT TTC TAC TGT 1089
      Met His Ser Phe Asn Cys Gln Gly Glu Phe Phe Tyr Cys
                                          360
                     355
      AAT ACA ACA AAG CTG TTT AAT AGT ACT TGG AAT GAT ACT 1128
      Asn Thr Thr Lys Leu Phe Asn Ser Thr Trp Asn Asp Thr
                                                  375
                              370
      ACA GAG TCA AAT AAC AAT GAT AGT ACT ATT ACA CTC CCA 1167
      Thr Glu Ser Asn Asn Asn Ser Thr Ile Thr Leu Pro
                                     385
                 380
30
      TGC AGA ATA AAA CAA ATT ATA AAC ATG TGG CAG GAA ATA 1206
      Cys Arg Ile Lys Gln Ile Ile Asn Met Trp Gln Glu Ile
                                             400
                         395
      GGA AAA GCA ATG TAT GCC CCT CCC ACC AGA GGA GAA ATT 1245
     Gly Lys Ala Met Tyr Ala Pro Pro Thr Arg Gly Glu Ile
35
                                                     415
                                 410
             405
      AAA TGT TCA TCA AAT ATT ACA GGA CTA CTG TTA ATA AGA 1284
      Lys Cys Ser Ser Asn Ile Thr Gly Leu Leu Leu Ile Arg
                                          425
                     420
     GAT GGT GGT ATT AAC ACT AGC GAT GCC ACC GAG ACC TTC 1323
40
      Asp Gly Gly Ile Asn Thr Ser Asp Ala Thr Glu Thr Phe
                             435
      AGA CCG GGA GGA GAT ATG AGG GAC AAT TGG AGA AGT 1362
      Arg Pro Gly Gly Gly Asp Met Arg Asp Asn Trp Arg Ser
                                     450
45
                 445
      GAA TTA TAT AAA TAT AAA GTA GTG AAA ATT GAG CCA TTA 1401
      Glu Leu Tyr Lys Tyr Lys Val Val Lys Ile Glu Pro Leu
                         460
                                             465
      455
      GGA GTA GCA CCC ACC AAG GCA AAG AGA AGA GTG GTG CAG 1440
      Gly Val Ala Pro Thr Lys Ala Lys Arg Arg Val Val Gln
50
                                 475
             470
      AGA GAA AAA AGA GCA GTA ACA CTA GGA GCT ATG TTC CTT 1479
      Arg Glu Lys Arg Ala Val Thr Leu Gly Ala Met Phe Leu
                      485
      GGG TTC TTA GGA GCA TAA AGC TTC 1503
55
      Gly Phe Leu Gly Ala Xaa Ser Phe
          495
                                  CLONE C6.5
          GGG GTA CCT GTA TGG AAA GAA GCA ACC ACC ACT CTA 36
60
          Gly Val Pro Val Trp Lys Glu Ala Thr Thr Thr Leu
                                               10
      TTT TGT GCA TCA GAT GCT AAA GCA TAT GAC ACA GAG GTG 75
      Phe Cys Ala Ser Asp Ala Lys Ala Tyr Asp Thr Glu Val
                                   20
65
```

	CAT	AAT	GTT	TGG	GCC	ACA	CAT	GCI	TGT	GTA	ccc	ACA	GAC	: 114
					30		His			35			_	
5			Pro				GTA Val 45						Glu	153
	TAD qe <i>A</i>	TTT Phe	AAC Asn	ATG Met 55	Trp	AAA Lys	AAT Asn	GAC Asp	ATG Met 60	GTA Val	GAA Glu	CAG	ATG	192
10	CAT His 65	GAG Glu	TNA Xaa	ATA	ATC	AGT Ser 70	TTA Leu	TGG Trp	GAT	CAA Gln	AGC Ser 75	CTA Leu	AAA Lys	231
	CCA	TGT	GTA Val	AAA	TTA	ACC	CCA Pro	CTC	TGT	ATT	ACT	TTA	AAT	270
15			80				AAT	85	_				90	
	Cys	Thr	Asn	Trp	Lys 95	Glu	Asn	Asp	Thr	Lys 100	Thr	Asn	Ser	
20	AGT Ser	AGT Ser 105	ACT Thr	ACA Thr	ACT Thr	AAT Asn	AAT Asn 110	AGT Ser	AGT Ser	GCT Ala	ACA Thr	GCT Ala 115	AAT Asn	348
	AGT Ser	AGT Ser	AGT Ser	ACT Thr 120	ACA Thr	ACT Thr	AAT Asn	AGT Ser	AGT Ser 125	TGG Trp	GGA Gly	GAG	ATA Ile	387
25				GAA			AAC Asn		TCT					426
30	ACA Thr	GGC	ATA Ile 145	AGA Arg	GAC Asp	AAG Lys	GTG Val	AAG Lys 150	AAA Lys	GAA Glu	TAT	GCA Ala	CTT Leu 155	465
							GTA Val						AAT	504
35							AGT Ser 175			ACC				543
							GTA Val					ATT		582
40 .	Ile 195	His	Tyr	TGT Cys	Thr	Pro 200	GCT Ala	Gly	TTT Phe	Ala	Ile 205	Leu	Lys	
45	Cys	Lys	Asp 210	Lys	Lys	Phe	AAT Asn	Gly 215	Thr	Gly	Pro	Cys	Lys 220	
	Asn	Val	Ser	Thr	Val 225	Gln	CAa	Thr	His	Gly 230	Ile	Lys	Pro	
50	Val	Val 235	Ser	Thr	Gln	Leu	CTG Leu 240	Leu	Asn	Gly	Ser	Leu 245	Ala	
	Glu	Glu	Glu	Val 250	Ile	Ile	AGA Arg	Ser	Ala 255	Asn	Phe	Ser	Asn	
55	Asn 260	Ala	Lys	Ile	Ile	Ile 265	GTA Val	Gln	Leu	Lys	Glu 27	Pro	Val	
60	GAA Glu	Ile	AAT Asn 275	TGT Cys	ACA Thr	AGA Arg	CCC Pro	AGC Ser 280	AAC . Asn .	AAT Asn	ACA Thr	Ile	AAA Lys 285	855
		ATA	CAC	Ile			GGG Gly	AGA	Ala :			GCA	ACA	894

```
GGA GAC ATA CGA GGA GAT ATA AGA CAA GCA CAT TGT AAC 933
      Gly Asp Ile Arg Gly Asp Ile Arg Gln Ala His Cys Asn
                              305
      ATT AGT GGA GCA AAA TGG AAT AAC ACT TTA AAG AAG GTA 972
      Ile Ser Gly Ala Lys Trp Asn Asn Thr Leu Lys Lys Val
 5
                  315
                                      320
      GTT ATA AAA TTA AAA GAA CAA TTT CCA AAT AAA ACA ATA 1011
      Val Ile Lys Leu Lys Glu Gln Phe Pro Asn Lys Thr Ile
      325
                          330
                                              335
      GTC TTT AAC CAT TCC TCA GGA GGG GAC CCA GAA ATT GTA 1050
10
      Val Phe Asn His Ser Ser Gly Gly Asp Pro Glu Ile Val
                                  345
              340
      ATG CAC AGT TTT AAT TGT CAA GGG GAA TTT TTC TAC TGT 1089
      Met His Ser Phe Asn Cys Gln Gly Glu Phe Phe Tyr Cys
                      355
                                          360
15
      AAT ACA ACG AAG CTG TTT AAT AGT ACT TGG AAT GAT ACT 1128
      Asn Thr Thr Lys Leu Phe Asn Ser Thr Trp Asn Asp Thr
                                                  375
                              370
      ACA GAG TCA AAT AAC AAT GAT AGT ACT ATT ACA CTC CCA 1167
      Thr Glu Ser Asn Asn Asn Asp Ser Thr Ile Thr Leu Pro
20
                                     385
                 380
      TGC AGA ATA AAA CAA ATT ATA AAC ATG TGG CAG GAA GTA 1206
      Cys Arg Ile Lys Gln Ile Ile Asn Met Trp Gln Glu Val
                          395
                                              400
      GGA AAA GCA ATG TAT GCC CCT CCC ATC AGA GGA GAA ATT 1245
25
      Gly Lys Ala Met Tyr Ala Pro Pro Ile Arg Gly Glu Ile
                                 410
              405
      AAA TGT TCA TCA AAT ATT ACA GGA CTA CTG TTA ACA AGA 1284
      Lys Cys Ser Ser Asn Ile Thr Gly Leu Leu Thr Arg
30
                     420
                                          425
      GAT GGT GGT ATT AAC ACT AGC GAT GCC ACC GAG ACC TTC 1323
      Asp Gly Gly Ile Asn Thr Ser Asp Ala Thr Glu Thr Phe
                              435
      AGA CCG GGA GGA GGA GAT ATG AGG GAC AAT TGG AGA AGT 1362
      Arg Pro Gly Gly Gly Asp Met Arg Asp Asn Trp Arg Ser
                                      450
                 445
      GAA TTA TAT AAA TAT AAA GTA GTG AAA ATT GAG CCA TTA 1401
      Glu Leu Tyr Lys Tyr Lys Val Val Lys Ile Glu Pro Leu
                         460
      455
      GGA GTA GCA CCC ACC AAG GCA AAG AGA AGA GTG GTG CAG 1440
40
      Gly Val Ala Pro Thr Lys Ala Lys Arg Arg Val Val Gln
                                 475
             470
      AGA GAA AAA AGA GCA GTA ACA CTA GGA GCT ATG TTC CTT 1479
      Arg Glu Lys Arg Ala Val Thr Leu Gly Ala Met Phe Leu
45
                      485
     GGG TTC TTG GGA GCA TAA AGC TTC 1503
      Gly Phe Leu Gly Ala Xaa Ser Phe
                              500 501
          495
                                   CLONE C8.3
50
          GTA CCT GTA TGG AAA GAA GCA ACC ACC ACT CTA TTT 37
          Val Pro Val Trp Lys Glu Ala Thr Thr Thr Leu Phe
                                               10
      TGT GCA TCA GAT GCT AAA GCA TAT GAT ACA GAG GTA CAT 76
      Cys Ala Ser Asp Ala Lys Ala Tyr Asp Thr Glu Val His
55
                                   20
              15
      AAT GTT TGG GCT ACA CAT GCC TGT GTA CCC ACA GAC CCC 115
      Asn Val Trp Ala Thr His Ala Cys Val Pro Thr Asp Pro
                                           35
                       30
      AAC CCA CAA GAA GTA GTA TTG GAA AAT GTA ACA GAA AAT 154
60
      Asn Pro Gln Glu Val Val Leu Glu Asn Val Thr Glu Asn
                               45
                                                   50
          40
      TTT AAC ATG TGG AAA AAT AAC ATG GTA GAA CAG ATG CAT 193
      Phe Asn Met Trp Lys Asn Asn Met Val Glu Gln Met His
65
                                       60
```

		Asp					TGG					Lys		
5	Cys	Val	Eya 80	Leu	Thr	Pro	CTC Leu	Cys 85	Val	Thr	Leu	Asn	Cys 90	
	Thr	Asn	Leu	Glu	Asn 95	Ala	AAT Asn	Asn	The	Glu 100	Asn	Ala	Asn	
10	Asn	Thr 105	Asn	Asn	Tyr	Thr	TTG Leu 110	Gly	Met	Glu	Arg	Gly 115	Glu	
15							AAT Asn			Thr				388
							TAT Tyr							427
20							AAT Asn							466
	Ile	Ser	Cys	Asn	Thr 160	Ser	GTC Val	Ile	Thr	Gln 165	Ala	Cys	Pro	
25	AAG Lys	GTA Val 170	TCC Ser	TTT Phe	GAG Glu	CTA Leu	ATT Ile 175	CCC Pro	ATA Ile	CAT His	TAT Tyr	TGT Cys 180	GCC Ala	544
30	CCG Pro	GCT Ala	GGT Gly	TTT Phe 185	GCG Ala	ATT Ile	CTA Leu	AAG Lys	TGT Cys 190	AAA Lys	Asp	AAG Lys	AAG Lys	583
							TGT Cys							622
35	Gln	Cys	Thr 210	His	Gly	Ile	AGA Arg	Pro 215	Val	Val	Ser	Thr	Gln 220	
							CTA Leu							700
40	Ile	Arg 235	Sen	Glu	Asn	Ile	ACA Thr 240	Asp	Asn	Ala	Lys	Thr 245	Ile	
45	Ile	Val	Gln	Leu 250	Asn	Glu	TCT	Ile	Val 255	Ile	Asn	Cys	Thr	
	Arg 260	Pro	Asn	Yàu	Asn	Thr 265	AGA Arg	Lys	Ser	Ile	Asn 270	Ile	Gly	
50	Pro	Gly	Arg 275	Ala	Phe	Tyr	ACA Thr	Thr 280	Gly	Asp	Ile	Ile	Gly 285	
	Asp	Ile	Arg	Gln	Ala 290	His	Cys	Asn	Leu	Ser 295	Lys	Thr	Gln	
55	Trp	Glu 300	Lys	Thr	Leu	Arg	Gln 305	Ile	Ala	Ile	Lys	Leu 310	Glu	
60	Glu	Lys	Phe	Lys 315	Asn	Lys	ACA Thr	Ile	Ala 320	Phe	Asn	Lys	Ser	
							ATT							1012

```
TGT GGA GGG GAA TTT TTC TAC TGT AAT ACA ACA AAA CTG 1051
      Cys Gly Gly Glu Ph Phe Tyr Cys Asn Thr Thr Lys Leu
              34Ô
                                 345
      TTT AAT AGT ACC TGG AAT TTA ACA CAA CCG TTT AGT AAT 1090
      Phe Asn Ser Thr Trp Asn Leu Thr Gln Pro Phe Ser Asn
 5
                                         360
                      355
      ACC GGG AAT CGT ACT GAA GAG TTA AAT ATT ACA CTC CCA 1129
      Thr Gly Asn Arg Thr Glu Glu Leu Asn Ile Thr Leu Pro
                             370
                                                 375
        365
      TGC AGA ATA AAA CAA ATC ATA AAC TTG TGG CAG GAA GTA 1168
10
      Cys Arg Ile Lys Gln Ile Ile Asn Leu Trp Gln Glu Val
                                     385
                 380
      GGC AAA GCA ATG TAT GCC CCT CCC ATC AGA GGA CAA ATT 1207
      Gly Lys Ala Met Tyr Ala Pro Pro Ile Arg Gly Gln Ile
                          395
                                             400
15
      AGA TGT TCA TCA AAT ATT ACA GGG CTA CTA TTA ACA AGA 1246
      Arg Cys Ser Ser Asn Ile Thr Gly Leu Leu Leu Thr Arg
                                 410
             405
      GAT GGT GGA AGT AAC ACC GGT GAC AAC AGG ACT GAG ACC 1285
      Asp Gly Gly Ser Asn Thr Gly Asp Asn Arg Thr Glu Thr
20
                                         425
                     420
      TTT AGA CCT GGA GGA GGA GAT ATG AGG GAC AAT TGG AGA 1324
      Phe Arg Pro Gly Gly Gly Asp Met Arg Asp Asn Trp Arg
                                                 440
                              435
         430
      AGT GAA TTA TAT AAA TAT AAA GTA GTA AGA ATT GAA CCA 1363
25
      Ser Glu Leu Tyr Lys Tyr Lys Val Val Arg Ile Glu Pro
                                     450
      TTA GGA GTA GCA CCC ACC CAG GCA AAG AGA AGA GTG GTG 1402
      Leu Gly Val Ala Pro Thr Gln Ala Lys Arg Arg Val Val
                                             465
30
      455
                         460
      CAA AGA GAA AAA AGA GCA GTG GGG ATA GGA GCT ATG TTC 1441
      Gln Arg Glu Lys Arg Ala Val Gly Ile Gly Ala Met Phe
                               475
             470
     CTT GGG TTC TTG GGA GAT AA 1461
     Leu Gly Phe Leu Gly Asp
35
                       485 486
                                  CLONE C8.6
         GTA CCT GTG TGG AAA GAA GCA ACC ACC ACT CTA TTT 37
         Val Pro Val Trp Lys Glu Ala Thr Thr Thr Leu Phe
40
                                              10
     TGT GCA TCA GAT GCT AAA GCA TAT GAT ACA GAG GTA CAT 76
     Cys Ala Ser Asp Ala Lys Ala Tyr Asp Thr Glu Val His
                                   20
              15
     AAT GTT TGG GCT ACA CAT GCC TGT GTA CCC ACA GAC CCC 115
45
     Asn Val Trp Ala Thr His Ala Cys Val Pro Thr Asp Pro
                                         35
                      30
      AAC CCA CAA GAA GTA GTA TTG GAA AAT GTA ACA GAA AAT 154
      Asn Pro Gln Glu Val Val Leu Glu Asn Val Thr Glu Asn
                              45
                                                  50
50
         40
      TTT AAC ATG TGG AAA AAT AAC ATG GTA GAA CAG ATG CAT 193
      Phe Asn Met Trp Lys Asn Asn Met Val Glu Gln Met His
                                       60
                  55
      GAG GAT ATA ATC AGT TTA TGG GAT CAA AGT CTA AAG CCA 232
     Glu Asp Ile Ile Ser Leu Trp Asp Gln Ser Leu Lys Pro
55
                          70
                                              75
      TGT GTA AAA TTA ACC CCA CTC TGT GTT ACT TTA AAT TGC 271
      Cys Val Lys Leu Thr Pro Leu Cys Val Thr Leu Asn Cys
                                  85
      ACT AAT TTG GAG AAT GCT AAT AAT ACC GAG AAT GCT AAT 310
60
      Thr Asn Leu Glu Asn Ala Asn Asn Thr Glu Asn Ala Asn
                                         100
                      95
      AAT ACC AAT AAT TAT ACC TTG GGG ATG GAG AGA GGT GAA 349
      Asn Thr Asn Asn Tyr Thr Leu Gly Met Glu Arg Gly Glu
                             110
65
```

	AG! Ar	A AAI	AA A	п Су	s Ser	TTC Phe	C AA:	r Are	e Th	r Thi	A AG	C TT	A AGA	N 388
	Chi			120		~ ~ ~ .			125	5				
5	ARI	i nnu	2 G1:	a AAA	Lys	GA	TA	r GC	A TTO	J TTT	r TA	LAA 1	A CTI	427
•	130)		, Ly.	o Lys	135	yı	. nic	ı Le	ı Pne			3 Leu	
			\ GTA	A CA	ATA			r AG1	P ACC		140 רמד י	י ראכיני	·	466
	Asp	Val	. Val	Glr	lle	Ast	Asr	. Ser	The	Asr	TV	. Acc	, LTG	400
			145	5				150)		_		155	
10	ATA	AGT	TGI	CAA 1	ACC	TCA	GTC	TTA :	AC	CAG	GCC	TG7	י ררא	505
	Ile	e Ser	Cys	a Asr	Thr	Ser	· Val	Ile	The			Cys	Pro	
	A A C	CT3	TCC		160		3 ma			165				
	Lva	Val	Ser	· Phe	GAG Glu	Dro	Tle	. CCC	ATE	CAT	TAT	TGT	GCC	544
15	-1-	170	,		. 014		175		, 116	nia	ıyı	180		
	CCG	GCT	GGT	TTI	. CCC	ATT	CTA	AAG	TGT	' AAA	GAT	AAG	DAG	583
	Pro	Ala	Gly	Phe	Ala	Ile	Leu	Lys	Cys	Lys	Asp	Lvs	Lvs	303
				185					190)	_	_	-	
20	TTC	AAT	GGA	ACA	GGA	CCA	TGT	' AAA	AAT	GTC	AGG	ACA	GTA	622
20	195	ASI	GIY	Inr	Gly	200	Cys	Lys	Asn	Val			Val	
			ACA	CAT	GGA	200	ACA	CCA	CTN	C.T.A	205			
	Gln	Cys	Thr	His	Gly	Ile	Ara	Pro	Val	Val	Sor	ACT	CAA	661
			210		•			215					220	
25	CTA	CTG	TTA	AAT	GGC	AGT	CTA	GCA	GAA	GAA	GAG	ATA	GTA	700
	Leu	Leu	Leu	Asn	Gly	Ser	Leu	Ala	Glu	Glu	Glu	Ile	Val	
					225					230				
	TID	AGA	TUT	GAA	AAT	ATC	ACA	GAC	AAT	GCT	AAA	ACC	ATA	739
30	116	235	Ser	Gru	Asn	116	240	Asp	ASN	Ala	Lys		Ile	
	ATA		CAG	CTA	AAT	GAA		ATA	GTG	ልጥጥ	דממ	245	202	770
	Ile	Val	Gln	Leu	Asn	Glu	Ser	Ile	Val	Ile	Asn	Cvs	Thr	775
				250					255			-		
2.5	AGA	CCC	AAT	AAC	AAC	ACA	AGA	AAA	AGT	ATA	AAT	ATA	GGA	817
35	Arg	Pro	Asn	Asn	Asn	Thr	Arg	Lys	Ser	Ile	Asn	Ile	Gly	
	260	ccc	B.C.B	CCN	Tre C	265	B C B	202			270			
	Pro	Glv	Ara	Ala	TTC Phe	Tur	The	The	GGA	GAC	ATA	ATA	GGA	856
		1	275			- , -	• • • •	280	Gry	vah	116	116	285	
40	GAT	ATA	AGA	CAA	GCA	CAT	TGT	AAC	CTT	AGT	AAA	ACA	CAA	895
	Asp	Ile	Arg	Gln	Ala	His	Cys	Asn	Leu	Ser	Lys	Thr	Gln	
					290					295				
	TGG	GAA	AAA	ACG	TTA	AGA	CAG	ATA	GCT	ATA	AAA	TTA	GAA	934
45	rrp	300	Lys	1112	Leu	arg	305	TIE	ATA	iie	Lys		Glu	
	GAA		TTT	AAG	AAT	AAA		ATA	GCC	ጥጥጥ	ידממ	310	TCC	077
	Glu	Lys	Phe	Lys	Asn	Lys	Thr	Ile	Ala	Phe	Asn	Lvs	Ser	7/3
				315					320			•		
E 0	TCA	GGA	GGG	GAC	CCA	GAA	ATT	GTA	ATG	CAC	AGT	TTT	AAT	1012
50	Ser	GIY	Gly	qeA	Pro	Glu	Ile	Val	Met	His	Ser	Phe	Asn	
	325 TGT	GGA	GGG	GC A	ጥጥጥ	330 TTC	ጥአሩ	ጥርመ			335			
	Cvs	Glv	Glv	Glv	TTT Phe	Pho	TVE	CA a	AGT Se-	ACG	AGA	AAA	CTG	1051
	-,-	,	340	1			- 7 -	345	Jer	THE	ur d	ràz	Leu J50	
55	TTT	AAT	AGT	ACC	TGG .	AAT	TTA	ACA	CAA	CCG	TTT	AGT	AAT	1090
	Phe	Asn	Ser	Thr	Trp	Asn	Leu	Thr	Gln	Pro	Phe	Ser	Asn	
					355					360				
	ACC	GGG	GAT	CGT	ACT	GAA	GAG	TTA	AAT	ATT	ACA	CTC	CCA	1129
60	inr	G1y 365	ASP	Arg	Thr			Leu	Asn	Ile			Pro	
00			ATA	222	CAA .		370 ATA	220	ጥጥ ~	TC	~~~	375		
	Cys	Ara	Ile	Lvš	Gln	Ile	Ile	Asn	i i G	IGG (Trn :	CAG (GAA	GTA :	1198
	•	- 3		380	- - · ·				385	p	GIII '	G 1 U	A G T	
									-					

```
GGC AAA GCA ATG TAT GCC CCT CCC ATC AGA GGA CAA ATT 1207
      Gly Lys Ala Met Tyr Ala Pro Pro Ile Arg Gly Gln Ile
                          395
                                                400
      AGA TGT TCA TCA AAT ATT ACA GGG CTA CTA TTA AGG AGA 1246
      Arg Cys Ser Ser Asn Ile Thr Gly Leu Leu Arg Arg
 5
                                  410
             405
      GAT GGT GGA AGT AAC ACC AGT GAC AAC CAG ACT GAG ACC
      Asp Gly Gly Ser Asn Thr Ser Asp Asn Gln Thr Glu Thr
                                           425
                      420
      TTT AGA CCT GGG GGA GGA GAT ATG AGG GAC AAG TGG AGA 1324
10
      Phe Arg Pro Gly Gly Gly Asp Met Arg Asp Lys Trp Arg
                                                   440
                              435
      AGT GAA TTA TAT AAA TAT AAA GTA GTA AGA ATT GAA CCA 1363
      Ser Glu Leu Tyr Lys Tyr Lys Val Val Arg Ile Glu Pro
                  445
                                       450
15
      TTA GGA GTA GCA CCC ACC CAG GCA AAG AGA AGA GTG GTG 1402
      Leu Gly Val Ala Pro Thr Gln Ala Lys Arg Arg Val Val
                                               465
                          460
      455
      CAA AGA GAA AAA AGA GCA GTG GGG ATA GGA GCT ATG TTC 1441
      Gln Arg Glu Lys Arg Ala Val Gly Ile Gly Ala Met Phe
20
                                  475
             470
      CTT AGG TTC TTA GGA GAT AAA GCT TCT AGA GTC 1474
      Leu Arg Phe Leu Gly Asp Lys Ala Ser Arg Val
                      485
25
                                   CLONE C15.2
          CTC GAG GTA CCT GTA TGG AAA GAA GCA ACT ACC ACT 36.
Leu Glu Val Pro Val Trp Lys Glu Ala Thr Thr
                                                 10
      CTA TTT TGT GCA TCA GAT GCT AAA GCA TAT AAT ACA GAG 75
30
      Leu Phe Cys Ala Ser Asp Ala Lys Ala Tyr Asn Thr Glu
                                                         25
                                    20
               15
      AAA CAT AAT GTT TGG GCC ACA CAC GCC TGT GTA CCC ACA 114
Lys His Asn Val Trp Ala Thr His Ala Cys Val Pro Thr
35
      GAT CCC AAC CCA CAA GAA GTA GTA TTG GGA AAT GTG ACA 153
      Asp Pro Asn Pro Gln Glu Val Val Leu Gly Asn Val Thr
                                                    50
                               45
      GAA AAT TTT AAC ATG TGG AAA AAT AAC ATG GTA GAA CAA 192
      Glu Asn Phe Asn Met Trp Lys Asn Asn Met Val Glu Gln
40
                                        60
                   55
      ATG CAT GAA GAT ATA ATC AGT TTA TGG GAT CAA AGT CTA 231
      Met His Glu Asp Ile Ile Ser Leu Trp Asp Gln Ser Leu
                                                75
                           70
      AAG CCA TGT GTA AAA TTA ACC CCA CTC TGT GTT ACT TTA 270
45
      Lys Pro Cys Val Lys Leu Thr Pro Leu Cys Val Thr Leu
                                    85
              80
      AAT TGC ACT GAT GAT TTA GGG AAT GCT ACT AAT ACC AAT 309
      Asn Cys Thr Asp Asp Leu Gly Asn Ala Thr Asn Thr Asn
                                           100
50
                       95
      AGT AGT GCC ACT ACC AAT AGT AGT AGT TGG GAA GAA ATG 348
      Ser Ser Ala Thr Thr Asn Ser Ser Ser Trp Glu Glu Met
          105
                               110
      AAG GGG GAA ATG AAA AGA TGC TCT TTC AAT ATC ACC ACA 387
      Lys Gly Glu Met Lys Arg Cys Ser Phe Asn Ile Thr Thr
55
                                       125
                  120
      AGC ATA AGA GAT AAG ATT AAG AAA GAA CAT GCA CTT TTC 426
      Ser Ile Arg Asp Lys Ile Lys Lys Glu His Ala Leu Phe
                          135
      130
      TAT AGA CTT GAT GTA GTA CCA ATA GAT AAT GAT AAT ACC 465
60
      Tyr Arg Leu Asp Val Val Pro Ile Asp Asn Asp Asn Thr
                                    150
               145
```

	AC Th	A TA	T Ac	G TT g Le	u Ile	e Ası	T TG	T AA S As	T AC	C TC.	A GT r Va	C AT	T ACA e Thr	504
					160	0				16	5			
5	GT	n Al 17	а Су О	s Pr	o Ly	s Vai	l Se:	r Pho 5	e Gl	u Pro	o Il	e Pro	C ATA	!
	CA:	r TT	T TG	T GC	C CCC	GC1	r GG	r TT	r GC	G AT	CT.			582
	HI	s Ph	е су	s Al 18	a Pro	o Ala	Gly	y Pho			e Le	u Lys	Cys	
10	AA'	r aa	T AA	GAC	G TTC	GAC	GGZ	A AA	190 A GG	A CC	ነ ጥርና	וממ ז	AAT	621
	ASI	n As	n Ly	s Th	r Phe	Glu	Gly	Ly	3 Gl	/ Pro	Cys	E Lys	. Asu	021
	19:	3				200)				209	5		
	Va l	. AG L Se	r AC	A GT/ r Va	A CAA	TGC	AC	A CAT	GG	ATT	AGO	CCA	GTA Val	660
15			21	0				219	5			-	220	
	GTC	TC	A AC	T CA	CTG	CTG	TTA	LAA A	GGC	AGT	CTA	A GCA	CAA	699
	Val	Se	r Th	r Gl	1 Leu 22	Leu	Leu	. Asr	ı Gly			a Ala	Glu	
	GAA	GA	G GT	A ATA	ATT	' AGA	тст	' GAC	144	23 ATC	. ,		AAT	730
20	· Glu	Gl	u Va	l Ile	Ile	Arg	Ser	Asp	Asr	Ile	Thr	Ast	Asn	/30
		23	5				240)				245		
	ACT	T.V	A ACC	C ATT	: ATA	GTA	CAG	CTA	AAC	GAA	TCT	GTA	GTA	777
				250)				255	ı				
25	ATT	'AA	r TG	C ACA	AGA	CCC	AAC	AAC	AAT	ACA	AGA	AAA	AGT	816
	260	ASI	ı Cys	Thr	Arg	Pro 265	Asn	Asn	Asn	Thr	Arg	Lys	Ser	
			TATA	GGA	CCA		AGT	GCA	ጥጥጥ	ттт	270	202	GGA	955
	Ile	His	Ile	Gly	Pro	Gly	Ser	Ala	Phe	Phe	Ala	Thr	GLV	022
30			275	i				280					285	
	GAA	ATA	ATA	GGA	GAT Asp	ATA	AGA	CAA	GCA	CAC	TGT	AAC	CTT	894
					290					295				
25	AGT	AGA	ACA	CAA	TGG	AAT	AAC	ACT	TTA	GGA	AAG	ATA	GTC	933
35	Ser	Arg	Thr	Gln	Trp	Asn	Asn	Thr	Leu	Gly	Lys	Ile	Val	
	ATA			AGA	GAA	CAA	305 TTT	AGA	444	CAA	TTTT	310	GAA	022
	Ile	Lys	Leu	Arg	Glu	Gln	Phe	Arg	Lys	Gln	Phe	Gly	Glu	312
40				315					320			-		
40	Lvs	Thr	Tle	GTC: Val	Phe	AAT	Ara	TCC	TCA	GGA	GGG	GAC	CCG	1011
	325					330					335			
	GAA	ATT	GCA	ATG	CAC	AGT	TTT	AAT	TGT	GGA	GGG	GAA	TTT	1050
45	GIU	ile	340	Met	His	Ser	Phe	Asn 345	Cys	Gly	Gly	Glu		
	TTC	TAC			ACA	ACA	GCA	CTG	ттт	AAT	AGT	ACC	350 TGG	1089
	Phe	Tyr	Суз	Aşn	Thr	Thr	Ala	Leu	Phe	Asn	Ser	Thr	Trp	1007
					355					360			•	
50	Asn	Val	Thr	Lvs	Gly	Len	AAT	AAC	ACT Th-	GAA	GGA	AAT	AGC	1128
		J 65					370				_	775		
	ACA	GGA	GAT	GAA	AAT	ATC	ATA	CTC	CCA	TGT	AGA	ATA	AAA	1167
	Thr	GIA	Asp	Glu 380	Asn	Ile	Ile	Leu		Суз	Arg	Ile	Lys	
55	CAA	ATT	ATA	AAC	ATG	TGG	CAG	GAA	385 GTA	GGA	444	GC2	ATG	1206
	Gin	Ile	Ile	Asn	Met	Trp	Gln	Glu	Val	Gly	Lys	Ala	Met	-200
	390					395					400			
	Tvr	Ala	Pro	Pro	Ile	AGT Ser	GUA	CAA	ATT	AGA	TGT	TCA	TCA	1245
60			405					410					415	
	AAC	ATT	ACA	GGG	CTG	CTA	CTA	ACA	AGA	GAT	GGT	CCT	ACT '	1284
	Asn	Ile	Thr	Gly	Leu	Leu	Leu	Thr	Arg	Asb	Gly	Gly	Ser	
					420					425				

```
AAG AAC GAG AGC ATC ACC ACC GAG GTC TTC AGA CCT GGA 1323
      Lys Asn Glu Ser Ile Thr Thr Glu Val Phe Arg Pro Gly
                               435
          430
      GGA GGA GAT ATG AGG GAC AAT TGG AGA AGT GAA TTA TAT 1362
      Gly Gly Asp Met Arg Asp Asn Trp Arg Ser Glu Leu Tyr
                                      450
                  445
      AAA TAT AAA GTA GTA AAA ATT GAA CCA TTA GGA GTA GCG 1401
      Lys Tyr Lys Val Val Lys Ile Glu Pro Leu Gly Val Ala
                                              465
                          460
      CCC ACC AAG GCA AAG AGA AGA GTG GTG CAG AGA GAA AAA 1440
10
      Pro Thr Lys Ala Lys Arg Arg Val Val Gln Arg Glu Lys
                                  475
                                                      480
      AGA GCA GTG GGA ACA ATA GGA GCT ATG TTC CTT GGG TTC 1479
      Arg Ala Val Gly Thr Ile Gly Ala Met Phe Leu Gly Phe
                      485
                                           490
15
      TTG GGA GCA TAA AGC TTC TAG AGT CGA CCT GCA 1512
      Leu Gly Ala Xaa Ser Phe Xaa Ser Arg Pro Ala
                              500
                                   CLONE C15.3
20
          CTC GAG GTA CCT GTG TGG AAA GAA GCA ACT ACC ACT 36
          Leu Glu Val Pro Val Trp Lys Glu Ala Thr Thr Thr
      CTA TTT TGT GCA TCA GAT GCT AAA GCA TAT AAT ACA GAG 75
      Leu Phe Cys Ala Ser Asp Ala Lys Ala Tyr Asn Thr Glu
25
                                   20
             15
      AAA CAT AAT GTT TGG GCC ACA CAC GCC TGT GTA CCC ACA 114
      Lys His Asn Val Trp Ala Thr His Ala Cys Val Pro Thr
                                            35
                       30
      GAT CCC AAC CCA CAA GAA GTA GTA TTG GGA AAT GTG ACA 153
30
      Asp Pro Asn Pro Gln Glu Val Val Leu Gly Asn Val Thr
                                                   50
                               45
      GAA AAT TTT AAC ATG TGG AAA AAT AAC ATG GTA GAA CAA 192
      Glu Asn Phe Asn Met Trp Lys Asn Asn Met Val Glu Gln
                                        60
35
                   55
      ATG CAT GAA GAT ATA ATC AGT TTA TGG GAT CAA AGT CTA 231
       Met His Glu Asp Ile Ile Ser Leu Trp Asp Gln Ser Leu
                                                75
                            70
      AAG CCA TGT GTA AAA TTA ACC CCA CTC TGT GTT ACT TTA 270
      Lys Pro Cys Val Lys Leu Thr Pro Leu Cys Val Thr Leu
40
                                   85
               8O:
      AAT TGC ACT GAT GAT TTA GGG AAT GCT ACT AAT ACC AAT 309
      Asn Cys Thr Asp Asp Leu Gly Asn Ala Thr Asn Thr Asn
                                          100
                       95
      AGC AGT GCC ACT ACC AAT AGT AGT AGT TGG GAA GAA ATG 348
45
      Ser Ser Ala Thr Thr Asn Ser Ser Ser Trp Glu Glu Met
         105
                              110
                                                   115
      AAG GGG GAA ATG AAA AGG TGC TCT TTC AAT ATC ACC ACA 387
      Lys Gly Glu Met Lys Arg Cys Ser Phe Asn Ile Thr Thr
                                       125
50
                  120
      AGC ATA AGA GAT AAG ATT AAG AAA GAA CAT GCA CTT TTC 426
      Ser Ile Arg Asp Lys Ile Lys Lys Glu His Ala Leu Phe
                                               140
                         135
      TAT AGA CTT GAT GTA GTA CCA ATA GAT AAT GAT AAT ACC 465
      Tyr Arg Leu Asp Val Val Pro Ile Asp Asn Asp Asn Thr
55
                                   150
      ACA TAT AGG TTG ATA AAT TGT AAT ACC TCA GTC ATT ACA 504
      Thr Tyr Arg Leu Ile Asn Cys Asn Thr Ser Val Ile Thr
                                          165
                      160
      CAG GCC TGT CCA AAG GTA TCA TTT GAG CCA ATT CCC ATA 543 Gln Ala Cys Pro Lys Val Ser Phe Glu Pro Ile Pro Ile
60
                              175
                                                   180
          170
      CAT TTT TGT GCC CCG GCT GGT TTT GCG ATT CTA AAG TGT 582
      His Ph Cys Ala Pro Ala Gly Phe Ala Ile L u Lys Cys
                                       190
65
                  185
```

```
AAT AAT AAG ACG TTC GAG GGA AAA GGA CCA TGT AAA AAT 621
       Asn Asn Lys Thr Phe Glu Gly Lys Gly Pro Cys Lys Asn
       195
                           200
                                               205
       GTC AGT ACA GTA CAA TGC ACA CAT GGA ATT AGG CCA GTA 660
  5
       Val Ser Thr Val Gln Cys Thr His Gly Ile Arg Pro Val
               210
                                   215
                                                       220
       GTG TCA ACT CAA CTG CTG TTA AAT GGC AGT CTA GCA GAA 699
       Val Ser Thr Gln Leu Leu Leu Asn Gly Ser Leu Ala Glu
                       225
                                           230
       GAA GAG GTA ATA ATT AGA TCT GGC AAT ATC ACA GAC AAT 738
 10
       Glu Glu Val Ile Ile Arg Ser Gly Asn Ile Thr Asp Asn
                               240
                                                   245
       ACT AAA ACC ATT ATA GTA CAG CTA AAC GAA TCT GTA GTA 777
       Thr Lys Thr Ile Ile Val Gln Leu Asn Glu Ser Val Val
 15
                  250
                                       255
       ATT AAT TGT ACA AGA TCC AAC AAC AAT ACA AGA AAA AGT 816
       Ile Asn Cys Thr Arg Ser Asn Asn Asn Thr Arg Lys Ser
                          265
                                              270
       ATA CAT ATA GGA CCA GGG AGT GCA TTT TTT GCA ACA GGA 855
20
       Ile His Ile Gly Pro Gly Ser Ala Phe Phe Ala Thr Gly
              275
                                   280
      GAA ATA ATA GGA GAT ATA AGA CAA GCA CAC TGT AAC CTT 894
      Glu Ile Ile Gly Asp Ile Arg Gln Ala His Cys Asn Leu
                      290
25
      AGT AGA ACA CAA TGG AAT AAC ACT TTA GGA AAG ATA GTC 933
      Ser Arg Thr Gln Trp Asn Asn Thr Leu Gly Lys Ile Val
          300
                              305
                                                   310
      ATA AAA TTA AGA GAA CAA TTT AGA AAA CAA TTT GGA GAA 972
      Ile Lys Leu Arg Glu Gln Phe Arg Lys Gln Phe Gly Glu
30
                 315
                                     320
      AAA ACA ATA GTC TTT AAT CGA TCC TCA GGA GGG GAC CCG 1011
      Lys Thr Ile Val Phe Asn Arg Ser Ser Gly Gly Asp Pro
                          3.30
                                               335
      GAA ATT GCA ATG CAC AGT TTT AAT TGT GGA GGG GAA TTT 1050
35
      Glu Ile Ala Met His Ser Phe Asn Cys Gly Gly Glu Phe
              340
                                  345
                                                       350
      TTC TAC TGT AAC ACA ACA GCA CTG TTT AAT AGT ACC TGG 1089
      Phe Tyr Cys Asn Thr Thr Ala Leu Phe Asn Ser Thr Trp
                      355
                                          360
      AAT GTT ACT AAA GGG TTG AAT AAC ACT GAA GGA AAT AGC 1128
40
      Asn Val Thr. Lys Gly Leu Asn Asn Thr Glu Gly Asn Ser
        365
                              370
                                                  375
      ACA GGG GAT GAA AAT ATC ATA CTC CCA TGT AGA ATA AAA 1167
      Thr Gly Asp Glu Asn Ile Ile Leu Pro Cys Arg Ile Lys
45
                  380
                                      385
      CAA ATT ATA AAC ATG TGG CAG GAA GTA GGA AAA GCA ATG 1206
      Gin Ile Ile Asn Met Trp Gin Glu Val Gly Lys Ala Met
      390
                         395
                                              400
      TAT GCC CCT CCC ATC AGT GGA CAA ATT AGA TGT TCA TCA 1245
50
      Tyr Ala Pro Pro Ile Ser Gly Gln Ile Arg Cys Ser Ser
             405
                                 410
                                                      415
      AAT ATT ACA GGG CTG CTA CTA ACA AGA GAT GGT GGT AGT 1284
      Asn Ile Thr Gly Leu Leu Thr Arg Asp Gly Gly Ser
                      420
                                          425
55
      AAG AAC GAG AGC ATC ACC ACC GAG GTC TTC AGA CCT GGA 1323
      Lys Asn Glu Ser Ile Thr Thr Glu Val Phe Arg Pro Gly
         430
                             435
                                                 440
      GGA GGA GAT ATG AGG GAC AAT TGG AGA AGT GAA TTA TAT 1362
      Gly Gly Asp Met Arg Asp Asn Trp Arg Ser Glu Leu Tyr
60
                                     450
     AAA TAT AAA GTA GTA AAA ATT GAA CCA TTA GGA GTA GCG 1401
     Lys Tyr Lys Val Val Lys Ile Glu Pro Leu Gly Val Ala
                         460
```

```
CCC ACC AAG GCA AAG AGA AGA GTG GTG CAG AGA GAA AAA 1440
      Pro Thr Lys Ala Lys Arg Arg Val Val Gln Arg Glu Lys
                                  475
      AGA GCA GTG GGA ACA ATA GGA GCT ATG TTC CTT GGG TTC 1479
     Arg Ala Val Gly Thr Ile Gly Ala Met Phe Leu Gly Phe
                      485
      TTA GGA GCA TAA AGC TTC TAG A 1501
      Leu Gly Ala Xaa Ser Phe Xaa
          495
10
                                    CLONE C7.2
     GG GAA TTC GGA TCC GGG GTA CCT GTG TGG AAG GAA GCA 38
         Glu Phe Gly Ser Gly Val Pro Val Trp Lys Glu Ala
     ACC ACC ACT CTA TTC TGT GCA TCA GAT GCT AGA GCA TAT 77
15
     Thr Thr Thr Leu Phe Cys Ala Ser Asp Ala Arg Ala Tyr
                                   20
              15
      GAC ACA GAG GTA CAT AAT GTT TGG GCC ACA CAT GCC TGT 116
     Asp Thr Glu Val His Asn Val Trp Ala Thr His Ala Cys
                       30
                                            35
20
     GTA CCC ACA GAC CCT AGT CCA CAA GAA GTA GTT TTG GAA 155
      Val Pro Thr Asp Pro Ser Pro Gln Glu Val Val Leu Glu
         40
                               45
                                                    50
      AAT GTG ACA GAA AAT TTT AAC ATG TGG AAA AAT AAC ATG 194
      Asn Val Thr Glu Asn Phe Asn Met Trp Lys Asn Asn Met
25
                                       60
                  55
      GTA GAA CAA ATG CAT GAG GAT ATA ATT AGT TTA TGG GAT 233
      Val Glu Gln Met His Glu Asp Ile Ile Ser Leu Trp Asp
                          70
     CAA AGC TTA AAG CCA TGT GTA AAA TTA ACC CCA CTC TGT 272
30
      Gln Ser Leu Lys Pro Cys Val Lys Leu Thr Pro Leu Cys
                                   85
              80
     GTT ACT TTA AAT TGC AGT GAT TAT AGG AAT GCT ACT GAT 311 Val Thr Leu Asn Cys Ser Asp Tyr Arg Asn Ala Thr Asp
                                          100
35
      TAT AAG AAT GCT ACT GAT ACC ACT AGT AGT AAC GAG GGA 350
     Tyr Lys Asn Ala Thr Asp Thr Thr Ser Ser Asn Glu Gly
                              110
         105
      AAG ATG GAG AGA GGA GAA ATA AAA AAC TGC TCT TTC AAT 389
     Lys Met Glu Arg Gly Glu Ile Lys Asn Cys Ser Phe Asn
                                       125
                 120
      ATT ACC ACA AGC ATA AAA AAT AAG ATG CAG AAA GAA TAT 428
      Ile Thr Thr Ser Ile Lys Asn Lys Met Gln Lys Glu Tyr
                         135
                                              140
      130
      GCA CTT TTC TAT AAA CTT GAT ATA GTA CCA ATA GAT AAT 467
45
      Ala Leu Phe Tyr Lys Leu Asp Ile Val Pro Ile Asp Asn
                                  150
              145
      ACA AGC TAT ACA TTG ATA AGT TGT AAC ACC TCA GTC ATT 506
      Thr Ser Tyr Thr Leu Ile Ser Cys Asn Thr Ser Val Ile
                                          165
                      160
50
      ACA CAG GCC TGT CCA AAG GTA TCC TTT GAA CCA ACT CCC 545
      Thr Gln Ala Cys Pro Lys Val Ser Phe Glu Pro Thr Pro
                              175
                                                   180
         170
     ATA CAT TAT TGT GCT CCG GCT GGT TTT GCG ATT CTA AAG 584
      Ile His Tyr Cys Ala Pro Ala Gly Phe Ala Ile Leu Lys
55
                                      190
                 185
      TGT AAT GAT AAG AAG TTC AGT GGA AAA GGA GAA TGT AAA 623
      Cys Asn Asp Lys Lys Phe Ser Gly Lys Gly Glu Cys Lys
                                               205
                          200
      AAT GTC AGC ACA GTA CAA TGT ACA CAT GGA ATT AGG CCA 662
60
      Asn Val Ser Thr Val Gln Cys Thr His Gly Ile Arg Pro
                                                       220
                                   215
              210
      GTA GTA TCA ACT CAA CTG CTG TTA AAT GGC AGT CTA GCA 701
      Val Val Ser Thr Gln Leu Leu Leu Asn Gly Ser Leu Ala
65
                      225
```

```
GAA GAA GAG GTG GTA ATT AGA TCT GAC AAT TTC ATA GAC 740
        Glu Glu Glu Val Val Ile Arg Ser Asp Asn Phe Ile Asp
                                24Õ
        AAT ACT AAA ACC ATA ATA GTA CAG CTG AAA GAA TCT GTA 779
        Asn Thr Lys Thr Ile Ile Val Gln Leu Lys Glu Ser Val
                    250
                                        255
       GAA ATT AAT TGT ATA AGA CCC AAC AAT AAT ACA AGA AAA 818
       Glu Ile Asn Cys Ile Arg Pro Asn Asn Asn Thr Arg Lys
                           265
                                               270
       GGT ATA CAT ATA GGA CCA GGG AGA GCA TGG TAT GCA ACA 857
 10
       Gly Ile His Ile Gly Pro Gly Arg Ala Trp Tyr Ala Thr
               275
                                    280
       GGA GAA ATA GTA GGA GAT ATA AGA AAG GCA TAT TGT AAC 896
       Gly Glu Ile Val Gly Asp Ile Arg Lys Ala Tyr Cys Asn
 15
                       290
                                            295
       ATT AGT AGA ACA AAA TGG AAT AAC ACT TTA ATA CAG ATA 935
       Ile Ser Arg Thr Lys Trp Asn Asn Thr Leu Ile Gln Ile
           300
                                305
                                                    310
       GCT AAC AAA TTA AAA GAA AAA TAT AAT ACA ACA ATA AGC 974
 20
       Ala Asn Lys Leu Lys Glu Lys Tyr Asn Thr Thr Ile Ser
                  315
                                        320
       TTT AAT CGA TCC TCA GGA GGG GAC CCA GAA ATT GTA ACG 1013
       Phe Asn Arg Ser Ser Gly Gly Asp Pro Glu Ile Val Thr
                           330
                                                335
 25
       CAT AGT TTT AAT TGT GGA GGG GAG TTT TTC TAC TGT GAT 1052
       His Ser Phe Asn Cys Gly Gly Glu Phe Phe Tyr Cys Asp
               340
                                    345
       TCA ACA CAA CTG TTT AAT AGT ACT TGG AAT TTA AAT GGT 1091
Ser Thr Gln Leu Phe Asn Ser Thr Trp Asn Leu Asn Gly
                       355
                                           360
       ACT TGG AAT TTT ACT GCA GGG TCA AAT GAA ACT GAA GGC 1130
       Thr Trp Asn Phe Thr Ala Gly Ser Asn Glu Thr Glu Gly
           365
                               370
                                                    375
      AAT ATC ACA CTC CCA TGC AGA ATA AAA CAA ATT ATA AAC 1169
      Asn Ile Thr Leu Pro Cys Arg Ile Lys Gln Ile Ile Asn
                  380
                                       385
      AGG TGG CAG GAA GTA GGG AAA GCA ATG TAT GCC CCT CCC 1208
      Arg Trp Gln Glu Val Gly Lys Ala Met Tyr Ala Pro Pro
      390
                          395
40
      ATC AGT GGA CAA ATA AAA TGC TCA TCA AAC ATT ACA GGG 1247
      Ile Ser Gly Gln Ile Lys Cys Ser Ser Asn Ile Thr Gly
              405
                                  410
      ATG ATA TTA ACA AGG GAT GGT GGT AAC GAG AAC AAT AAT 1286
      Met Ile Leu Thr Arg Asp Gly Gly Asn Glu Asn Asn Asn
45
                      420
                                          425
      GAG AGC AGT ACT ACT GAG ACC TTC AGA CCG GGA GGA GGA 1325
      Glu Ser Ser Thr Thr Glu Thr Phe Arg Pro Gly Gly Gly
         430
                              435
      GAT ATG AGG AAC AAT TGG AGA AGT GAA TTA TAT AAA TAT 1364
      Asp Met Arg Asn Asn Trp Arg Ser Glu Leu Tyr Lys Tyr
                  445
                                      450
      AAA GTA GTA AAA ATT GAA CCA TTA GGA GTA GCA CCC ACC 1403
      Lys Val Val Lys Ile Glu Pro Leu Gly Val Ala Pro Thr
                          460
                                               465
55
      AAG GCA AAG AGA AGA GTG GTG CAG AGA GAA AAA AGA GCA 1442
      Lys Ala Lys Arg Arg Val Val Gln Arg Glu Lys Arg Ala
             470
                                   475
      GTG GGA GCG CTA GGA GCT ATG TTC CTT GGG TTC TTA GGA 1481
      Val Gly Ala Leu Gly Ala Met Phe Leu Gly Phe Leu Gly
60
                      485
                                          490
      GCA TAA AGC TTC TAG ACC GAC TCT AGA GGA TCC 1514
      Ala Xaa Ser Phe Xaa Thr Asp Ser Arg Gly Ser
          495
                              500
```

```
CLONE C7.10
          GTA CCT GTG TGG AAG GAA GCA ACC ACC CTA TTC 37 Val Pro Val Trp Lys Glu Ala Thr Thr Leu Phe
                                                10
      TGT GCA TCA GAT GCT AGA GCA TAT GAC ACA GAG GTA CAT 76
 5
      Cys Ala Ser Asp Ala Arg Ala Tyr Asp Thr Glu Val His
                                    20
               15
      AAT GTT TGG GCC ACA CAT GCC TGT GTA CCC ACA GAC CCT 115
      Asn Val Trp Ala Thr His Ala Cys Val Pro Thr Asp Pro
10
                       30
      AGT CCA CAA GAA GTA TTT TTG GGA AAT GTG ACA GAA AAT 154
      Ser Pro Gln Glu Val Phe Leu Gly Asn Val Thr Glu Asn
                                45
      TTT AAT ATG TGG AAA AAT AAC ATG GTA GAA CAA ATG TAT 193
15
      Phe Asn Met Trp Lys Asn Asn Met Val Glu Gln Met Tyr
                   55
                                       60
      GAG GAT ATA ATT AGT TTA TGG GAT CAA AGC TTA AAG CCA 232
      Glu Asp Ile Ile Ser Leu Trp Asp Gln Ser Leu Lys Pro
                           70
                                                75
      TGT GTA AAA TTA ACC CCA CTC TGT GTT ACT TTA AAT TGC 271
20
      Cys Val Lys Leu Thr Pro Leu Cys Val Thr Leu Asn Cys
               80
                                   85
      AGT GAT TAT AGG AAT GCT ACT GAT TAT AAG AAT GCT ACT 310
      Ser Asp Tyr Arg Asn Ala Thr Asp Tyr Lys Asn Ala Thr
25
                       95
                                           100
      GAT ACC ACT AGT AGT AAC GAG GGA AAG ATG GAG AGA GGA 349
      Asp Thr Thr Ser Ser Asn Glu Gly Lys Met Glu Arg Gly
                              110
      GAA ATA AAA AAC TGC TCT TTC AAT ATC ACC ACA AGC ATA 388
      Glu Ile Lys Asn Cys Ser Phe Asn Ile Thr Thr Ser Ile
                                      125
                  120
      AAA AAT AAG ATG CAG AAA GAA TAT GCA CTT TTC TAT AAA 427
      Lys Asn Lys Met Gln Lys Glu Tyr Ala Leu Phe Tyr Lys
                         135
      CTT AAT ATA GTA CCA ATA GAT AAT ACA AGC TAT ACA TTG 466
      Leu Asn Ile Val Pro Ile Asp Asn Thr Ser Tyr Thr Leu
             145
                                  150
                                                       155
      ATA AGT TGT AAC ACC TCA GTC ATT ACA CAG GCC TGT CCA 505
      Ile Ser Cys Asn Thr Ser Val Ile Thr Gln Ala Cys Pro
40
                      160
                                           165
      AAG GTA TCC TTT GAA CCA ATT CCC ATA CAT TAT TGT GCT 544
      Lys Val Ser Phe Glu Pro Ile Pro Ile His Tyr Cys Ala
                              175
         170
                                                   180
      CCG GCT GGT TTT GCG ATT CTA AAG TGT AAT GAT AAG AAG 583
45
      Pro Ala Gly Phe Ala Ile Leu Lys Cys Asn Asp Lys Lys
                                      190
                  185
      TTC AGT GGA AAA GGA GAA TGT AAA AAT GTC AGC ACA GTA 622
      Phe Ser Gly Lys Gly Glu Cys Lys Asn Val Ser Thr Val
                         200
                                              205
      195
50
      CAA TGT ACA CAT GGA ATT AGG CCA GTA GTA TCA ACT CAA 661
      Gln Cys Thr His Gly Ile Arg Pro Val Val Ser Thr Gln
                                  215
      CTG CTG TTA AAT GGC AGT CTA GCA GAA GAA GAG GTG GTA 700
      Leu Leu Leu Asn Gly Ser Leu Ala Glu Glu Val Val
55
                     225
                                         230
      ATT AGA TOT GAC AAT TTO ACA GAC AAT ACT AAA ACC ATA 739
      Ile Arg Ser Asp Asn Phe Thr Asp Asn Thr Lys Thr Ile
                              240
      ATA GTA CAG CTG AAA GAA TCT GTA GAA ATT AAT TGT ATA 778
      Ile Val Gln Leu Lys Glu Ser Val Glu Ile Asn Cys Ile
60
                                     255
                  250
      AGA CCC AAC AAT AAT ACA AGA AAA GGT ATA CAT ATA GGA 817
      Arg Pro Asn Asn Asn Thr Arg Lys Gly Ile His Ile Gly
                          265
```

```
CCA GGG AGA GCA TGG TAT GCA ACA GGA GAA ATA GTA GGA 856
        Pro Gly Arg Ala Trp Tyr Ala Thr Gly Glu Ile Val Gly
                275
                                     280
        GAT ATA AGA CAG GCA TAT TGT AAC ATT AGT AGA ACA AAA 895
        Asp Ile Arg Gln Ala Tyr Cys Asn Ile Ser Arg Thr Lys
                        290
                                             295
       TGG AAT AAC ACT TTA ATA CAG ATA GCT AAC AAA TTA AAA 934
       Trp Asn Asn Thr Leu Ile Gln Ile Ala Asn Lys Leu Lys
                                305
                                                    310
       GAA AAA TAT AAT ACA ACA ATA AGC TTT AAT CGA TCC TCA 973
 10
       Glu Lys Tyr Asn Thr Thr Ile Ser Phe Asn Arg Ser Ser
                   315
                                        320
       GGA GGG GAC CCA GAA ATT GTA ACC CAT AGT TTT AAT TGT 1012
       Gly Gly Asp Pro Glu Ile Val Thr His Ser Phe Asn Cys
 15
                            330
                                                335
       GGA GGG GAA TTT TTC TAC TGT AAT TCA ACA CAA CTG TTT 1051
       Gly Gly Glu Phe Phe Tyr Cys Asn Ser Thr Gln Leu Phe
               340
                                    345
                                                        350
       AAT AGT ACT TGG AAT TTA AAT GGT ACT TGG AAT TTT ACT 1090 Asn Ser Thr Trp Asn Leu Asn Gly Thr Trp Asn Phe Thr
 20
                       355
                                           360
       GCA GGG TCA AAT GAA ACT GAA GGC AAT ATC ACA CTC CCA 1129
       Ala Gly Ser Asn Glu Thr Glu Gly Asn Ile Thr Leu Pro
           365
                                370
                                                    375
       TGC AGA ATA AAA CAA ATT ATA AAC AGG TGG CAG GAA GTA 1158
 25
       Cys Arg Ile Lys Gln Ile Ile Asn Arg Trp Gln Glu Val
                   380
                                       385
       GGA AAA GCA ATG TAT GCC CCT CCC ATC AGT GGA CAA ATA 1207 Gly Lys Ala Met Tyr Ala Pro Pro Ile Ser Gly Gln Ile
30
       390
                           395
                                               400
       AGA TGC TCA TCA AAC ATT ACA GGG ATG ATA TTA ACA AGG 1246
      Arg Cys Ser Ser Asn Ile Thr Gly Met Ile Leu Thr Arg
              405
                                  410
      GAT GGT GGT AAC GAG AAC AAT AAT GAG AGC AGT ACT ACT 1285
35
      Asp Gly Gly Asn Glu Asn Asn Glu Ser Ser Thr Thr
                      420
                                           425
      GAG ACC TTC AGA CCG GGA GGA GGA GAT ATG AGG AAC AAT 1324
      Glu Thr Phe Arg Pro Gly Gly Gly Asp Met Arg Asn Asn
         430
                               435
      40
      Trp Arg Ser Glu Leu Tyr Lys Tyr Lys Val Val Lys Ile
                  445
                                       450
      GAG CCA TTA GGA GTA GCA CCC ACC GAC TCT AGA GGA TCC 1402
      Glu Pro Leu Gly Val Ala Pro Thr Asp Ser Arg Gly Ser
45
                           460
      TCT AGA 1408
      Ser Arg
          469
50
                                    CLONE C11.5
          GAG GTA CCT GTG TGG AAA GAA GCA ACC ACT ACT CTA 36
          Glu Val Pro Val Trp Lys Glu Ala Thr Thr Thr Leu
      TTT TGT GCA TCA GAT GCT AAA GCA TAT GAC ACA GGG GTG 75
55
      Phe Cys Ala Ser Asp Ala Lys Ala Tyr Asp Thr Gly Val
                                   20
                                                        25
      CAT AAT GTT TGG GCC ACA CAT GCC TGT GTA CCC ACA GAC 114
      His Asn Val Trp Ala Thr His Ala Cys Val Pro Thr Asp
60
      CCC AAC CCA CAA GAA ATA GAA TTG GTA AAT GTG ACA GAA 153
      Pro Asn Pro Gln Glu Ile Glu Leu Val Asn Val Thr Glu
                               45
                                                   50
     GAT TTT AAC ATG TGG AAA AAT AAA ATG GTA GAC CAG ATG 192
     Asp Phe Asn Met Trp Lys Asn Lys Met Val Asp Gln Met
65
                   55
```

```
CAT GAG GAT ATA ATC AGT TTA TGG GAT GAA AGC CTA AAG 231
      His Glu Asp Ile Ile Ser Leu Trp Asp Glu Ser Leu Lys
                           70
      CCA TGT GTA AAG TTA ACC CCA CTT TGT GTT ACT CTA AAC 270
      Pro Cys Val Lys Leu Thr Pro Leu Cys Val Thr Leu Asn
               80
                                   85
      TGC AGT GAT GTG AAC AAT TCC ACA AAT CCT AAT GAT ACT 309
      Cys Ser Asp Val Asn Asn Ser Thr Asn Pro Asn Asp Thr
                                          100
                      95
      AAT ACT AAT TCC ACT AAT ACT ACT TCC TCT ACT CCT ACG 348
10
      Asn Thr Asn Ser Thr Asn Thr Thr Ser Ser Thr Pro Thr
                                                   115
                              110
      GCC ACT ACT AGT AGC GAG GAA AAG ATG GAG AAG GGA GAA 387
      Ala Thr Thr Ser Ser Glu Glu Lys Met Glu Lys Gly Glu
15
                  120
                                      125
     ATA AAA AAC TGC TCT TTC AAT ATC ACC ACA CAC ATG AAA 426 Ile Lys Asn Cys Ser Phe Asn Ile Thr Thr His Met Lys
                          135
                                               140
      GAT AAG GCA CAG AAA GAA TAT GCA CTT TTT TAT AAA CTT 465
      Asp Lys Ala Gln Lys Glu Tyr Ala Leu Phe Tyr Lys Leu
20
                                  150
      GAT ATA GTA CCA ATA GAT GAT AAT AAT GCC AGC TAT AGG 504
      Asp Ile Val Pro Ile Asp Asp Asn Asn Ala Ser Tyr Arg
                                          165
                      160
      TTG ATA AGT TGT AAT ACC TCA GAC ATT ACA CAG GCC TGT 543
25
      Leu Ile Ser Cys Asn Thr Ser Asp Ile Thr Gln Ala Cys
                                                 180
                              175
        170
      CCA AAG GTG ACC TTT GAG CCA ATT CCC ATA CAT TAT TGT 582
      Pro Lys Val Thr Phe Glu Pro Ile Pro Ile His Tyr Cys
                                      190
30
                  185
      GCC CCG GCT GGT TTT GCG ATT CTA AAG TGT AAA GAT AAG 621
      Ala Pro Ala Gly Phe Ala Ile Leu Lys Cys Lys Asp Lys
                          200
                                              205
      195
      AAG TTC AAT GGA ACA GGA CCA TGT TCA AAG GTC AGC ACA 660
      Lys Phe Asn Gly Thr Gly Pro Cys Ser Lys Val Ser Thr
35
            210
                                  215
                                                       220
      GTA CAA TGT ACA CAT GGA ATT AGG CCA GTA GTA TCA ACT 699
      Val Gln Cys Thr His Gly Ile Arg Pro Val Val Ser Thr
                                          230
                      225
      CAA CTG TTG TTA AAT GGC AGT CTT GCA GAA GAA GAA GTA 738
40
      Gln Leu Leu Asn Gly Ser Leu Ala Glu Glu Glu Val
                                                  245
                              240
      GTA ATT AGA TCT GTC AAT TTC ACA GAC AAT GCT AAA ATC 777
      Val Ile Arg Ser Val Asn Phe Thr Asp Asn Ala Lys Ile
                  250
                                      255
45
      ATA ATA GTA CAG CTG AAA GAA CCT GTA GCA ATT AAT TGT 816
      Ile Ile Val Gln Leu Lys Glu Pro Val Ala Ile Asn Cys
                        265
                                              270
      ACA AGA CCC AAC AAC AAT ACA AGA AAA GGT ATA CAT CTA 855
      Thr Arg Pro Asn Asn Asn Thr Arg Lys Gly Ile His Leu
50
                                  280
              275
      GGA CCA GGG AGC ACA TTT TAT ACA ACA GGA GAA ATA ATA 894
      Gly Pro Gly Ser Thr Phe Tyr Thr Thr Gly Glu Ile Ile
                                          295
                      290
      GGA GAC ATA AGA AAA GCA TAT TGC AAG ATT AGT AAA GAA 933
      Gly Asp Ile Arg Lys Ala Tyr Cys Lys Ile Ser Lys Glu
                              305
         300
      AAA TGG AAT AAC ACT TTA AGA CAG GTA GTT AAA AAA TTA 972
      Lys Trp Asn Asn Thr Leu Arg Gln Val Val Lys Lys Leu
                                      320
60
                  315
      AGA GAA CAA TTT GGG AAT AAA ACA ATA ATT TTT AAT CGA 1011
      Arg Glu Gln Phe Gly Asn Lys Thr Ile Ile Phe Asn Arg
```

```
TCC TCA GGA GGG GAC CCA GAA ATT GTA ATG CAC AGT TTT 1050
        Ser Ser Gly Gly Asp Pro Glu Ile Val Met His Ser Phe
               340
                                    345
        AAC TGT GGA GGG GAG TTT TTC TAC TGT AAT ACA ACA CAA 1089
   5
       Asn Cys Gly Glu Phe Phe Tyr Cys Asn Thr Thr Gln
                       355
                                           360
       CTG TTT AAT AGT ACT TGG AAT AAT ACT GAA GGG ACA AAT 1128
       Leu Phe Asn Ser Thr Trp Asn Asn Thr Glu Gly Thr Asn
                               370
                                                   375
       AGC ACT GAA GGA AAT AGC ACA ATC ACA CTC CCA TGC AGA 1167
 10
       Ser Thr Glu Gly Asn Ser Thr Ile Thr Leu Pro Cys Arg
                  380
                                       385
       ATA AAA CAA ATT ATA AAT ATG TGG CAG GAA GTA GGA AAA 1206
       Ile Lys Gln Ile Ile Asn Met Trp Gln Glu Val Gly Lys
 15
                           395
                                               400
       GCA ACG TAT GCC CCT CCC ATC AGA GGA CGA ATT AGA TGC 1245
       Ala Thr Tyr Ala Pro Pro Ile Arg Gly Arg Ile Arg Cys
               405
                                  410
                                                      415
       ATA TCA AAT ATT ACA GGA CTG CTA TTA ACA AGA GAT GGT 1284
       Ile Ser Asn Ile Thr Gly Leu Leu Leu Thr Arg Asp Gly
 20
                      420
                                          425
       GGT AGG AAT GTC ACA AAC AAT ACC GAA ACC TTC AGA CCT 1323
       Gly Arg Asn Val Thr Asn Asn Thr Glu Thr Phe Arg Pro
                              435
                                                  440
      GGA GGA GGA GAC ATG AGG GAC AAT TGG AGA AGT GAA TTA 1362
 25
       Gly Gly Gly Asp Met Arg Asp Asn Trp Arg Ser Glu Leu
                  445
                                      450
      TAT AAA TAT AAA GTA GTA AAA GTT GAA CCA TTA GGA ATA 1401
       Tyr Lys Tyr Lys Val Val Lys Val Glu Pro Leu Gly Ile
 30
                          460
                                              465
      GCA CCC ACC AAG GCA AAG AGA AGA GTG GTG CAC AGA GAC 1440
      Ala Pro Thr Lys Ala Lys Arg Arg Val Val His Arg Asp
                                  475
      AAA AGA GCA GCA CTA GGA GCC TTG TTC CTT GGG TTC TTA 1479
35
      Lys Arg Ala Ala Leu Gly Ala Leu Phe Leu Gly Phe Leu
                      485
      GGA GCA TAA AAG CTT CTA GA 1499
      Gly Ala Xaa Lys Leu Leu
           495
40
                                  CLONE C11.7
          GAG GTA CCT GTA TGG AAA GAA GCA ACC ACT ACT CTA 36
          Glu Val Pro Val Trp Lys Glu Ala Thr Thr Leu
                                               10
45
      TTT TGT GCA TCA GAT GCT AAA GCA TAT GAC ACA GAG GTG 75
      Phe Cys Ala Ser Asp Ala Lys Ala Tyr Asp Thr Glu Val
              15
                                  20
      CAT AAT GTT TGG GCC ACA CAT GCC TGT GTA CCC ACA GAC 114
      His Asn Val Trp Ala Thr His Ala Cys Val Pro Thr Asp
50
                      30
                                          35
      CCC AAC CCA CAA GAA ATA GAA TTG GTA AAT GTG ACA GAA 153
      Pro Asn Pro Gln Glu Ile Glu Leu Val Asn Val Thr Glu
         40
                              45
      GAT TTT AAC ATG TGG AAA AAT AAA ATG GTA GAC CAG ATG 192
55
      Asp Phe Asn Met Trp Lys Asn Lys Met Val Asp Gln Met
                  55
                                      60
      CAT GAG GAT ATA ATC AGT TTA TGG GAT GAA AGC CTA AAG 231
     His Glu Asp Ile Ile Ser Leu Trp Asp Glu Ser Leu Lys
                          70
                                              75
      CCA TGT GTA AAG TTA ACC CCA CTT TGT GTT ACT CTA AAC 270
60
     Pro Cys Val Lys Leu Thr Pro Leu Cys Val Thr Leu Asn
              80
                                  85
     TGC AGT GAT GTG AAC AAT TCC ACA AAT CCT AAT GAT ACT 309
     Cys Ser Asp Val Asn Asn Ser Thr Asn Pro Asn Asp Thr
65
                                         100
```

```
AAT ACT AAT TCC ACT AAT ACT ACT TCC TCT ACT CCT ACG 348
       Asn Thr Asn Ser Thr Asn Thr Thr Ser Ser Thr Pro Thr
                                 110
       GCC ACT ACT AGT AGC GAG GAA AAG ATG GAG AAG GGA GAA 387
  5
       Ala Thr Thr Ser Ser Glu Glu Lys Met Glu Lys Gly Glu
                    120
                                          125
       ATA AAA AAC TGC TCT TTC AAT ATC ACC ACA CAC ATG AAA 426
       Ile Lys Asn Cys Ser Phe Asn Ile Thr Thr His Met Lys
                            135
                                                   140
       GAT AAG GTA CAG AAA GAA TAT GCA CTT TTT TAT AAA CTT 465
10
       Asp Lys Val Gln Lys Glu Tyr Ala Leu Phe Tyr Lys Leu
               145
                                     150
       GAT ATA GTA CCA ATA GAT GAT AAT AAT ACC AGC TAT AGG 504
       Asp Ile Val Pro Ile Asp Asp Asn Asn Thr Ser Tyr Arg
15
                        160
                                              165
       TTG ATA AGT TGT AAT ACC TCA GTC ATT ACA CAG GCC TGT 543
       Leu Ile Ser Cys Asn Thr Ser Val Ile Thr Gln Ala Cys
          170
                                 175
       CCA ATG GTG ACC TTT GAG CCA ATT CCC ATA CAT TAT TGT 582
Pro Met Val Thr Phe Glu Pro Ile Pro Ile His Tyr Cys
20
                   185
                                         190
      GCC CCG GCT GGT TTT GCG ATT CTA AAG TGT AAA GAT AAG 621
Ala Pro Ala Gly Phe Ala Ile Leu Lys Cys Lys Asp Lys
                            200
      AAG TTC AAT GGA ACA GGA CCA TGT TCA AAG GTC AGC ACA 660
25
      Lys Phe Asn Gly Thr Gly Pro Cys Ser Lys Val Ser Thr
                                    215
               210
      GTA CAA TGT ACA CAT GGA ATT AGG CCA GTA GTA TCA ACT 699
      Val Gln Cys Thr His Gly Ile Arg Pro Val Val Ser Thr
30
                        225
                                              230
      CAA CTG TTG TTA AAT GGC AGT CTT GCA GAA GAA GAA GTA 738
      Gln Leu Leu Leu Asn Gly Ser Leu Ala Glu Glu Val
                                 240
      GTA ATT AGA TCT GTC AAT TTC ACA GAC AAT GCT AAA ATC 777
35
      Val Ile Arg Ser Val Asn Phe Thr Asp Asn Ala Lys Ile
                                         255
                   250
      ATA ATA GTA CAG CTG AAA GAA CCT GTA GCA ATT AAT TGT 816
      Ile Ile Val Gln Leu Lys Glu Pro Val Ala Ile Asn Cys
      260
                            265
                                                  270
      ACA AGA CCC AAC AAC AAT ACA AGA AAA GGT ATA CAT CTA 855
40
      Thr Arg Pro Asn Asn Asn Thr Arg Lys Gly Ile His Leu
      275 280 285
GGA CCA GGG AGC ACA TTT TAT ACA ACA GGA GAA ATA ATA 894
      Gly Pro Gly Ser Thr Phe Tyr Thr Thr Gly Glu Ile Ile
45
                       290
                                             295
      GGA GAC ATA AGA AAA GCA TAT TGC AAG ATT AGT AAA GAA 933
Gly Asp Ile Arg Lys Ala Tyr Cys Lys Ile Ser Lys Glu
          300
                                305
                                                     310
      AAA TGG AAT AAC ACT TTA AGA CAG GTA GTT AAA AAA TTA 972
Lys Trp Asn Asn Thr Leu Arg Gln Val Val Lys Lys Leu
50
                                        320
                  315
      AGA GAA CAA TTT GGG AAT AAA ACA ATA ATT TTT AAT CGA 1011
      Arg Glu Gln Phe Gly Asn Lys Thr Ile Ile Phe Asn Arg
      325
                            330
                                                  335
55
      TCC TCA GGA GGG GAC CCA GAA ATT GTA ATG CAC AGT TTT 1050
      Ser Ser Gly Gly Asp Pro Glu Ile Val Met His Ser Phe
               340
                                     345
      AAC TGT GGA GGG GAG TTT TTC TAC TGT AAT ACA ACA CAA 1089
      Asn Cys Gly Glu Phe Phe Tyr Cys Asn Thr Thr Gln
60
                       355
                                             360
      CTG TTT AAT AGT ACT TGG AAT AAT ACT GAA GGG ACA AAT 1128
      Leu Phe Asn Ser Thr Trp Asn Asn Thr Glu Gly Thr Asn
                                370
```

```
AGC ACT GAA GGA AAT AGC ACA ATC ACA CTC CCA TGC AGA 1167
       S r Thr Glu Gly Asn Ser Thr Ile Thr Leu Pro Cys Arg
                   380
                                       385
       ATA AAA CAA ATT ATA AAT ATG TGG CAG GAA GTA GGA AAA 1206
       Ile Lys Gln Ile Ile Asn Met Trp Gln Glu Val Gly Lys
   5
       390
                           395
       GCA ACG TAT GCC CCT CCC ATC AGA GGA CGA ATT AGA TGC 1245
       Ala Thr Tyr Ala Pro Pro Ile Arg Gly Arg Ile Arg Cys
               405
                                  410
       ATA TCA AAT ATT ACA GGA CTG CTA TTA ACA AGA GAT GGT 1284
 10
       Ile Ser Asn Ile Thr Gly Leu Leu Leu Thr Arg Asp Gly
                       420
                                           425
       GGT AGG AAT GTC ACA AAC AAT ACC GAN NCC TTC AGA CCT 1323
       Gly Arg Asn Val Thr Asn Asn Thr Xaa Xaa Phe Arg Pro
 15
          430
                              435
                                                 440
       GGA GGA GGA GAC ATG AGG GAC AAT TGG AGA AGT GAA TTA 1362
       Gly Gly Gly Asp Met Arg Asp Asn Trp Arg Ser Glu Leu
                                      450
       TAT AAA TAT AAA GTA GTA AAA GTT GAA CCA TTA GGA ATA 1401
       Tyr Lys Tyr Lys Val Val Lys Val Glu Pro Leu Gly Ile
 20
                          460
       GCA CCC ACC AAG GCA AAG AGA AGA GTG GTG CAC AGA GAC 1440
       Ala Pro Thr Lys Ala Lys Arg Arg Val Val His Arg Asp
                                  475
                                                      480
       AAA AGA GCA GCA CTA GGA GCT TTG TTC CTT GGG TTC TTA 1479
 25
       Lys Arg Ala Ala Leu Gly Ala Leu Phe Leu Gly Phe Leu
                       485
       GGA GCA TAA AAG CTT CTA GA 1499
       Gly Ala Xaa Lys Leu Leu
           495
                                  CLONE C10.5
      G GTA CCT GTG TGG AAA GAA GCA AAC ACA ACT CTA TTT 37
          Val Pro Val Trp Lys Glu Ala Asn Thr Thr Leu Phe
35
                            5
      TGT GCA TCA GAT GCT AAA GCA TAT GAT AGA GAA GTA CAT 76
      Cys Ala Ser Asp Ala Lys Ala Tyr Asp Arg Glu Val His
               15
                                    20
      AAT GTT TGG GCA ACA CAT GCC TGT GTA CCC ACA GAC CCC 115
40
      Asn Val Trp Ala Thr His Ala Cys Val Pro Thr Asp Pro
                       30
      AAC CCA CAA GAA ATA GTA TTG GGA AAT GTG ACA GAA AAT 154
      Asn Pro Gln Glu Ile Val Leu Gly Asn Val Thr Glu Asn
          40
                               45
                                                   50
      TTT AAC ATG TGG AAA AAT AAC ATG GTA GAA CAA ATG CAT 193
45
      Phe Asn Met Trp Lys Asn Asn Met Val Glu Gln Met His
                   5 5
                                      60
      GAG GAT ATA ATC AAT TTA TGG GAT CAA AGC TTA AAG CCA 232
      Glu Asp Ile Ile Asn Leu Trp Asp Gln Ser Leu Lys Pro
50
      65
                           70
                                               75
      TGT GTA AAG TTA ACT CCA CTC TGT GTT ACT TTA AAG TGC 271
      Cys Val Lys Leu Thr Pro Leu Cys Val Thr Leu Lys Cys
               80
                                   85
      AAG GAT CTG GAG AGG AAT ACT ACC TAT AAT AGC ACT ATT 310
55
      Lys Asp Leu Glu Arg Asn Thr Thr Tyr Asn Ser Thr Ile
                      95
                                         100
     ACC AAT AAT AGT AGT TTG GAG GGA CTA AGA GAA CAA ATG 349
      Thr Asn Asn Ser Ser Leu Glu Gly Leu Arg Glu Gln Met
                             110
      ACA AAC TGC TCT TTC AAC ATC ACC ACA AGT ATA AGA GAT 388
      Thr Asn Cys Ser Phe Asn Ile Thr Thr Ser Ile Arg Asp
                 120
                                      125
     AAG GTG CAG AAA GAA TAT GCA CTT TTG TAT AAA CTT GAT 427
     Lys Val Gin Lys Glu Tyr Ala Leu Leu Tyr Lys Leu Asp
65
```

```
GTA GTA CCA ATA GAA GAA GAT GAC AAT ACT AGC TAT AGA 466
      Val Val Pro Il Glu Glu Asp Asp Asn Thr Ser Tyr Arg
                                   150
                                                        155
              145
      TTG ATA AGT TGT AAC ACC TCA GTC ATT ACA CAG GCT TGT 505
      Leu Ile Ser Cys Asn Thr Ser Val Ile Thr Gln Ala Cys
                                           165
                       160
      CCA AAG ACA TCC TTT GAG CCA ATT CCC ATA CAT TAT TGT 544
      Pro Lys Thr Ser Phe Glu Pro Ile Pro Ile His Tyr Cys
                              175
         170
                                                   180
      GCC CCG GCT GGT TTT GCG ATT CTA AAG TGT AAT GAT AAG 583 Ala Pro Ala Gly Phe Ala Ile Leu Lys Cys Asn Asp Lys
10
                                       190
                  185
      AAG TTC AAT GGA ACA GGA CCA TGT AAA AAT GTC AGC ACA 622
      Lys Phe Asn Gly Thr Gly Pro Cys Lys Asn Val Ser Thr
                          200
      195
15
      GTA CAA TGT ACA CAT GGA ATT AGG CCA GTA GTA TCA ACT 661
      Val Gln Cys Thr His Gly Ile Arg Pro Val Val Ser Thr
              210
                                   215
      CAA CTG TTG TTA AAT GGC AGT CTA GCA GAA GAA GAG GTA 700
20
      Gln Leu Leu Asn Gly Ser Leu Ala Glu Glu Glu Val
                      225
                                          230
      GTA ATC AGA TCT GCC AAT TTC ACA GAC AAT GCT AAA ACC 739
      Val Ile Arg Ser Ala Asn Phe Thr Asp Asn Ala Lys Thr
                              240
                                                   245
      ATA ATA GTA CAT CTA AAT GAA ACT GTA AAA ATT AAT TGT 778
25
      Ile Ile Val His Leu Asn Glu Thr Val Lys Ile Asn Cys
                  250
                                       255
      ACA AGA CTT GGC AAC AAT ACA AGA AAA AGT ATA AAT ATA 817
      Thr Arg Leu Gly Asn Asn Thr Arg Lys Ser Ile Asn Ile
30
      260
                          265
                                              270
      GGA CCA GGG AGA GTA CTC TAT GCA ACA GGA GAA ATA ATA 856
      Gly Pro Gly Arg Val Leu Tyr Ala Thr Gly Glu Ile Ile
                                   280
              275
      GGA GAC ATA AGA CAA GCA CAT TGT AAC ATT AGT AGA GCA 895
      Gly Asp Ile Arg Gln Ala His Cys Asn Ile Ser Arg Ala
35
                      290
                                          295
      CAA TGG AAT AAG ACT TTA GAA AAG GTA GTT GAC AAA TTA 934
      Gln Trp Asn Lys Thr Leu Glu Lys Val Val Asp Lys Leu
                              305
      AGA AAA CAA TTT GGG GAT AAT ACA ACA ATA GCT TTT AAT 973
40
      Arg Lys Gln. Phe Gly Asp Asn Thr Thr Ile Ala Phe Asn
                                      320
                 315
      CGA TCC TCA GGA GGG GAC CCA GAA ATT GTA ATG CAC ACT 1012
      Arg Ser Ser Gly Gly Asp Pro Glu Ile Val Met His Thr
45
                                               335
      325
                          330
      TTT AAT TGT GGA GGG GAA TTT TTC TAC TGT AAT ACA ACA 1051
      Phe Asn Cys Gly Gly Glu Phe Phe Tyr Cys Asn Thr Thr
                                 345
      CAA CTG TTT AAT AGT ACT TGG AAT AAT ACT TGG AAG GAT 1090
      Gln Leu Phe Asn Ser Thr Trp Asn Asn Thr Trp Lys Asp
50
                      355
                                           360
      CCT AAC AGG AGT GAC AAT ATC ACA CTC CCA TGC AGA ATA 1129
      Pro Asn Arg Ser Asp Asn Ile Thr Leu Pro Cys Arg Ile
                               370
          365
                                                   375
      AAA CAA ATT ATA AAC ATG TGG CAG GAA GTA GGA AAA GCA 1168
55
      Lys Gln Ile Ile Asn Met Trp Gln Glu Val Gly Lys Ala
                                      385
                  380
      ATG TAC GCC CCT CCC ATC AGA GGG GAA ATT AGA TGT TCA 1207
      Met Tyr Ala Pro Pro Ile Arg Gly Glu Ile Arg Cys Ser
                                              400
60
      390
                          395
     TCA AAT ATC ACA GGG CTG CTA CTA ACA AGA GAT GGT GGT 1246
      Ser Asn Ile Thr Gly Leu Leu Leu Thr Arg Asp Gly Gly
                                   410
              405
```

```
AAT GAC GAT GGT AAT GAC ACG ACC ACA AAC AGG ACC GAG 1285
        Asn Asp Asp Gly Asn Asp Thr Thr Thr Asn Arg Thr Glu
                        420
                                            425
        ATC TTC AGA CCT GGA GGA GGA GAT ATG AGG GAC AAT TGG 1324
       Ile Phe Arg Pro Gly Gly Gly Asp Met Arg Asp Asn Trp
                                435
                                                    440
       Arg Ser Glu Leu Tyr Arg Tyr Lys Val Val Lys Ile Glu
                   445
                                        450
       CCA TTA GGA ATA GCA CCC ACC AGG GCA AAG AGA AGA GTG 1402
 10
       Pro Leu Gly Ile Ala Pro Thr Arg Ala Lys Arg Arg Val
                           460
                                                465
       GTG CAG AGA GAA AAA AGA GCA GTA GGA CTA GGA GCT TTG 1441
       Val Gln Arg Glu Lys Arg Ala Val Gly Leu Gly Ala Leu
 15
               470
                                   475
       TTC CTT GGG T TCTTAGGAG CATAAAGCTT CTAGA 1475
       Phe Leu Gly
 20
                                    CLONE C10.7
           GTA CCT GTG TGG AAA GAA GCA AAC ACA ACT CTA TTT 37
           Val Pro Val Trp Lys Glu Ala Asn Thr Thr Leu Phe
       TGT GCA TCA GAT GCT AAA GCA TAT GAT AGA GAA GTA CAT 76
 25
       Cys Ala Ser Asp Ala Lys Ala Tyr Asp Arg Glu Val His
                15
       AAT GTT TGG GCA ACA CAT GCC TGT GTA CCC ACA GAC CCC 115
       Asn Val Trp Ala Thr His Ala Cys Val Pro Thr Asp Pro
                        30
                                             35
       AAC CCA CAA GAA ATA GTA TTG GGA AAT GTG ACA GAA AAT 154
 30
       Asn Pro Gln Glu Ile Val Leu Gly Asn Val Thr Glu Asn
           40
                                45
                                                    50
       TTT AAC ATG TGG AAA AAT AAC ATG GTA GAA CAA ATG CAT 193
       Phe Asn Met Trp Lys Asn Asn Met Val Glu Gln Met His
35
                                        60
      GAG GAT ATA ATC AAT TTA TGG GAT CAA AGC TTA AAG CCA 232
      Glu Asp Ile Ile Asn Leu Trp Asp Gln Ser Leu Lys Pro
                            70
      TGT GTA AAG TTA ACT CCA CTC TGT GTT ACT TTA AAG TGC 271
40
      Cys Val Lys Leu Thr Pro Leu Cys Val Thr Leu Lys Cys
               80
                                   85
      AAG GAT CTG GAG AGG AAT ACT ACC TAT AAT AGC ACT ATT 310 Lys Asp Leu Glu Arg Asn Thr Thr Tyr Asn Ser Thr Ile
                        95
45
      ACC AAT AAT AGT AGT TTG GAG GGA CTA AGA GAA CAA ATG 349
      Thr Asn Asn Ser Ser Leu Glu Gly Leu Arg Glu Gln Met
         105
                             110
      ACA AAC TGC TCT TTC AAC ATC ACC ACA AGT ATA AGA GAT 388
Thr Asn Cys Ser Phe Asn Ile Thr Thr Ser Ile Arg Asp
50
                  120
                                       125
      AAG GTG CAG AAA GAA TAT GCA CTT TTG TAT AAA CTT GAT 427
      Lys Val Gln Lys Glu Tyr Ala Leu Leu Tyr Lys Leu Asp
      130
                         135
                                               140
      GTA GTA CCA ATA GAA GAA GAT GAC AAT ACT AGC TAT AGA 466
55
      Val Val Pro Ile Glu Glu Asp Asp Asn Thr Ser Tyr Arg
              145
                                  150
                                                       155
      TTG ATA AGT TGT AAC ACC TCA GTC ATT ACA CAG GCT TGT 505
      Leu Ile Ser Cys Asn Thr Ser Val Ile Thr Gln Ala Cys
                      160
                                           165
      CCA AAG ACA TCC TTT GAG CCA ATT CCC ATA CAT TAT TGT 544
60
      Pro Lys Thr Ser Phe Glu Pro Ile Pro Ile His Tyr Cys
          170
                              175
                                                  180
      GCC CCG GCT GGT TTT GCG ATT CTA AAG TGT AAT GAT AAG 583
      Ala Pro Ala Gly Phe Ala Ile Leu Lys Cys Asn Asp Lys
65
                                      190
```

```
AAG TTC AAT GGA ACA GGA CCA TGT AAA AAT GTC AGC ACA 622
      Lys Phe Asn Gly Thr Gly Pro Cys Lys Asn Val Ser Thr
                         200
      GTA CAA TGT ACA CAT GGA ATT AGG CCA GTA GTA TCA ACT 661
      Val Gln Cys Thr His Gly Ile Arg Pro Val Val Ser Thr
                                215
             210
      CAA CTG TTG TTA AAT GGC AGT CTA GCA GAA GAA GAG GTA 700
      Gln Leu Leu Asn Gly Ser Leu Ala Glu Glu Glu Val
                     225
                                         230
      GTA ATC AGA TCT GCC AAT TTC ACA GAC AAT GCT AAA ACC 739
10
     Val Ile Arg Ser Ala Asn Phe Thr Asp Asn Ala Lys Thr
                             240
                                                 245
      ATA ATA GTA CAT CTA AAT GAA ACT GTA AAA ATT AAT TGT 778
      Ile Ile Val His Leu Asn Glu Thr Val Lys Ile Asn Cys
                                    255
15
                250
     ACA AGA CTT GGC AAC AAT ACA AGA AAA AGT ATA AAT ATA 817
     Thr Arg Leu Gly Asn Asn Thr Arg Lys Ser Ile Asn Ile
                                             270
     260
                         265
     GGA CCA GGG AGA GTA CTC TAT GCA ACA GGA GAA ATA ATA 856
     Gly Pro Gly Arg Val Leu Tyr Ala Thr Gly Glu Ile Ile
20
                                280
                                                    285
            275
     GGA GAC ATA AGA CAA GCA CAT TGT AAC ATT AGT AGA GCA 895
     Gly Asp Ile Arg Gln Ala His Cys Asn Ile Ser Arg Ala
                    290
                                        295
     CAA TGG AAT AAG ACT TTA GAA AAG GTA GTT GAC AAG TTA 934
25
     Gln Trp Asn Lys Thr Leu Glu Lys Val Val Asp Lys Leu
                           305
     AGA AAA CAA TTT GGG GAT AAT ACA ACA ATA GCT TTT AAT 973
     Arg Lys Gln Phe Gly Asp Asn Thr Thr Ile Ala Phe Asn
30
                                     320
                 315
     CGA TCC TCA GGA GGG GAC CCA GAA ATT GTA ATG CAC ACT 1012
     Arg Ser Ser Gly Gly Asp Pro Glu Ile Val Met His Thr
                        330
                                            335
     TTT AAT TGT GGA GGG GAA TTT TTC TAC TGT AAT ACA ACA 1051
     Phe Asn Cys Gly Gly Glu Phe Phe Tyr Cys Asn Thr Thr
35
            340
                                 345
     CAA CTG TTT AAT AGT ACT TGG AAT AAT ACT TGG AAG GAT 1090
     Gln Leu Phe Asn Ser Thr Trp Asn Asn Thr Trp Lys Asp
                     355
                                         360
     CCT AAC AGG AGT GAC AAT ATC ACA CTC CCA TGC AGA ATA 1129
40
     Pro Asn Arg Ser Asp Asn Ile Thr Leu Pro Cys Arg Ile
        365
                             370
     AAA CAA ATT ATA AAC ATG TGG CAG GAA GTA GGA AAA GCA 1168
     Lys Gln Ile Ile Asn Met Trp Gln Glu Val Gly Lys Ala
                                    385
45
                 380
     ATG TAC GCC CCT CCC ATC AGA GGG GAA ATT AGA TGT TCA 1207
     Met Tyr Ala Pro Pro Ile Arg Gly Glu Ile Arg Cys Ser
                        395
                                            400
     390
     TCA AAT ATC ACA GGG CTG CTA CTA ACA AGA GAT GGT GGT 1246
50
     Ser Asn Ile Thr Gly Leu Leu Leu Thr Arg Asp Gly Gly
                                                    415
             405
                                410
     AAT GAC GAT GGT AAT GAC ACG ACC ACA AAC AGG ACC GAG 1285
     Asn Asp Asp Gly Asn Asp Thr Thr Thr Asn Arg Thr Glu
                                       425
                    420
     ATC TTC AGA CCT GGA GGA GGA GAT ATG AGG GAC AAT TGG 1324
55
     Ile Phe Arg Pro Gly Gly Gly Asp Met Arg Asp Asn Trp
                            435
     Arg Ser Glu Leu Tyr Arg Tyr Lys Val Val Lys Ile Glu
                                    450
60
                 445
     CCA TTA GGA ATA GCA CCC ACC AGG GCA AAG AGA AGA GTG 1402
     Pro Leu Gly Ile Ala Pro Thr Arg Ala Lys Arg Arg Val
                        460
```

```
GTG CAG AGA GAA AAA AGA GCA GTA GGA CTA GGA GCT TTG 1441
        Val Gln Arg Glu Lys Arg Ala Val Gly Leu Gly Ala Leu
                470
                                     475
        TTC CTT GGG TTC TTG GGA GCA TAA AGC TTC TAG A 1475
       Phe Leu Gly Phe Leu Gly Ala Xaa Ser Phe Xaa
                         485
                                     CLONE C17.1
            CTC GAG GTA CCT GTG TGG AAA GAA GCA ACC ACC ACT 36
            Leu Glu Val Pro Val Trp Lys Glu Ala Thr Thr Thr
 10
                             5
                                                10
       CTA TTT TGT GCA TCA GAT GCT AAA GCA TAT GAT TCA GAG 75
       Leu Phe Cys Ala Ser Asp Ala Lys Ala Tyr Asp Ser Glu
                 15
                                      20
       GCA CAT AAT GTT TGG GCC ACA CAT GCC TGT GTA CCC ACA 114 Ala His Asn Val Trp Ala Thr His Ala Cys Val Pro Thr
 15
                        30
       GAC CCC AAC CCA CAA GAA GTA GAA TTG GAA AAT GTG ACA 153 Asp Pro Asn Pro Gln Glu Val Glu Leu Glu Asn Val Thr
 20
           40
                                  45
                                                       50
       GAA AAT TTT AAC ATG TGG AAA AAT AAC ATG GTA GAA CAG 192 Glu Asn Phe Asn Met Trp Lys Asn Asn Met Val Glu Gln
                    55
                                         60
       ATG CAT GGG GAT ATA ATT AGT TTA TGG GAT CAA AGC CTA 231
       Met His Gly Asp Ile Ile Ser Leu Trp Asp Gln Ser Leu
                             70
       AAG CCA TGT GTA AAA TTA ACC CCA CTC TGT GTT ACG TTA 270
       Lys Pro Cys Val Lys Leu Thr Pro Leu Cys Val Thr Leu
               80
                                     85
       AAT TGC ACT GAC CCA AAT GTT ACT AAT AGC GAG AGA ACG 309
30
      Asn Cys Thr Asp Pro Asn Val Thr Asn Ser Glu Arg Thr
                         95
                                            100
       ATA GAG GGG GGA GAA ATA AAA AAT TGC TCT TTC AAT ATC 348
       Ile Glu Gly Gly Glu Ile Lys Asn Cys Ser Phe Asn Ile
35
         105
                               110
                                                     115
      ACC ACA AAC ATA AGA GAT AGG TTT CAG AAA GAA TAT GCA 387
      Thr Thr Asn Ile Arg Asp Arg Phe Gln Lys Glu Tyr Ala
                   120
                                        125
      CTT TTT TAT AAA CTT GAT GTA ATA CCA TTA GGT AAT GAT 426
      Leu Phe Tyr Lys Leu Asp Val Ile Pro Leu Gly Asn Asp
40
                         135
      AAT ACT AGC TAT AGG TTG ATA AGT TGT AAC ACC TCA GTC 465
      Asn Thr Ser Tyr Arg Leu Ile Ser Cys Asn Thr Ser Val
              145
                                    150
                                                         155
45
      ATT ACA CAG GCC TGT CCA AAG GTA TCC TTT GAG CCA ATT 504
      Ile Thr Gln Ala Cys Pro Lys Val Ser Phe Glu Pro Ile
                      160
                                             165
      CCC ATA CAT TAT TGT GCC CCG GCT GGT TTT GCG ATT CTA 543
      Pro Ile His Tyr Cys Ala Pro Ala Gly Phe Ala Ile Leu
50
          170
                                175
                                                    180
      AAG TGT AAA GAT AAG AAG TTC AAT GGA ACA GGA CCA TGT 582
      Lys Cys Lys Asp Lys Lys Phe Asn Gly Thr Gly Pro Cys
                  185
                                        190
      ACA AAT GTC AGC ACA GTA CAA TGT ACA CAT GGA ATT AAG 621
55
      Thr Asn Val Ser Thr Val Gln Cys Thr His Gly Ile Lys
                           200
                                                205
      CCA GTA GTA TCA ACT CAA CTG TTG TTA AAT GGC AGT CTA 660
      Pro Val Val Ser Thr Gln Leu Leu Leu Asn Gly Ser Leu
                                    215
60
      GCA GAA GAA GAC ATA GTA ATT AGA TCC GCC AAT CTC ACA 699
      Ala Glu Glu Asp Ile Val Ile Arg Ser Ala Asn Leu Thr
                       225
      GAC AAT GCT AAA AAC ATA ATA GTA CAG CTG AAT GAA TCT 738
      Asp Asn Ala Lys Asn Ile Ile Val Gln Leu Asn Glu Ser
65
                               240
```

```
GTA ACA ATG AAT TGT ACA AGA CCC AAC AAC AAT ACA ATG 777
       Val Thr Met Asn Cys Thr Arg Pro Asn Asn Asn Thr Met
                    250
                                         255
       AAA AGT ATA CAT ATA GGA CCA GGC AGA GCA TTT TAT GCA 816
  5
       Lys Ser Ile His Ile Gly Pro Gly Arg Ala Phe Tyr Ala
       260
                            265
                                                 270
       ACA GGA AAC ATA ATA GGA GAT ATA AGA CAA GCA CAT TGT 855
       Thr Gly Asn Ile Ile Gly Asp Ile Arg Gln Ala His Cys
               275
                                    280
       AAC ATT AGT GGA ACA AAA TGG AAT GAC ACT TTG AAA AAG 894 Asn Ile Ser Gly Thr Lys Trp Asn Asp Thr Leu Lys Lys
 10
                        290
                                             295
       ATA GCT ATA AAA TTA AGA GAA CAA TTT AAT AAG ACA ATA 933
       Ile Ala Ile Lys Leu Arg Glu Gln Phe Asn Lys Thr Ile
15
                               305
       GTC TTT AAT CAA TCC TCA GGA GGG GAC CCA GAA ATT GCA 972
Val Phe Asn Gln Ser Ser Gly Gly Asp Pro Glu Ile Ala
                   315
                                         320
       ACG CTC AGT TTT AAT TGT GGA GGG GAA TTT TTC TAC TGT 1011
20
       Thr Leu Ser Phe Asn Cys Gly Gly Glu Phe Phe Tyr Cys
                           330
                                                 335
       AAT TCA ACA CAA CTG TTT AAT AGT ACT TGG AAT AGT ACT 1050
      Asn Ser Thr Gln Leu Phe Asn Ser Thr Trp Asn Ser Thr
               340
                                    345
                                                         350
25
      GGG TCA AAT AAC ACT AAA GGA AAT GAC ACA ATC ACA CTC 1089
      Gly Ser Asn Asn Thr Lys Gly Asn Asp Thr Ile Thr Leu
                       355
                                            360
      CCA TGC AGA ATA AGA CAA ATT ATA AAC ATG TGG CAG AAA 1128
      Pro Cys Arg Ile Arg Gln Ile Ile Asn Met Trp Gln Lys
30
          365
                                370
                                                     375
      ATA GGA AAA GCA ATG TAT GCC CCT CCC ATC AAA GGG CAA 1167
      Ile Gly Lys Ala Met Tyr Ala Pro Pro Ile Lys Gly Gln
                   380
                                        385
      ATT AGA TGT TCA TCA AAT ATT ACA GGG CTA ATA TTA ACA 1206
      Ile Arg Cys Ser Ser Asn Ile Thr Gly Leu Ile Leu Thr
35
      390
                           395
                                                400
      AGA GAT GGT GGT AAC AAC AAC ATG AGC AAG ACC ACC GAG 1245
      Arg Asp Gly Gly Asn Asn Asn Met Ser Lys Thr Thr Glu
               405
                                    410
                                                        415
      ACC TTC AGA CCT GGA GGA GGA GAT ATG AGG GAC AAT TGG 1284
40
      Thr Phe Arg. Pro Gly Gly Gly Asp Met Arg Asp Asn Trp
                      420
                                           425
      45
                               435
      CCA TTA GGA GTA GCA CCC ACC AGG GCA AAG AGA AGA GTG 1362
      Pro Leu Gly Val Ala Pro Thr Arg Ala Lys Arg Arg Val
                   445
                                        450
      GTG CAG AGA GAA AAA AGA GCA GTG GGA ATA GGA GCT GTG 1401
50
      Val Gln Arg Glu Lys Arg Ala Val Gly Ile Gly Ala Val
                           460
                                                465
      TTC CTT GGG TTC TTG GGA GCA TAA AGC TTC TAG A 1435
      Phe Leu Gly Phe Leu Gly Ala Xaa Ser Phe Xaa
              470
55
                                    CLONE C17.3
          CTC GAG GTA CCT GTG TGG AAA GAA GCA ACC ACC ACT 36
          Leu Glu Val Pro Val Trp Lys Glu Ala Thr Thr Thr
60
      CTA TTT TGT GCA TCA GAT GCT AAA GCA TAT GAT TCA GAG 75
      Leu Phe Cys Ala Ser Asp Ala Lys Ala Tyr Asp Ser Glu
               15
                                     20
      GCA CAT AAT GTT TGG GCC ACA CAT GCC TGT GTA CCC ACA 114
Ala His Asn Val Trp Ala Thr His Ala Cys Val Pro Thr
65
```

	vah	40	ASD	Pro	GIN	GIU	ı va. 49	l Glu 5	ı Leu	Glu	Asn	Va:	The	
5	GIU	nen	Pne	A90 55	Met	Trp	Ly	3 Asr	Asn 60	Met	Val	GAA Glu	CAC Glr	
	65	ura	GLY	Asp	Ile	11e) Sei	. Leu	TGG	GAT Asp	Gln	Ser	Leu	
10	Lys	PIG	80	vai	Lys	TTA	ACC	Pro 85	Leu	Cys	GTT Val	ACC Thr	Leu	
15	ASII	cya	inr	Asp	Pro 95	Asn	Val	ACT Thr	AAT Asn	Ser	Glu	Arg	ACG Thr	309
	rre	105	GIÀ	GIA	Glu	Ile	Lys 110	AAT Asn	Cys	Ser	Phe	Asn	Ile	
20	Thr	Thr	Asn	11e 120	Arg	ysb	Arg	TTT Phe	Gln 125	Lys	Glu	TAT Tyr	GCA Ala	
2.5	130	File	Tyr	Lys	Leu	135	Val	ATA Ile	Pro	Leu	Gly 140	Asn	Asp	
25	ASN	Thr	Ser 145	Tyr	Arg	Leu	Ile	AGT Ser 150	Cys	Asn	ACC Thr	Ser	Val	
30	iie	Ing	GIN	Ala	Cys 160	Pro	Lys	GTA Val	Ser	Phe	Glu	Pro	Ile	
	Pro	170	HIS	Tyr	Cys	Ala	Pro 175	GCT Ala	Gly	Phe	Ala	Ile	Leu	
35	rys	Cys	Lys	Asp 185	Lys	Lys	Phe	TAA neA	Gly 190	Thr	Gly	Pro	Cys	
40	195	Asn	Val .	Ser	Thr	Val 200	Gln	TGT Cys	Thr	His (Gly 205	Ile	Lys	
40	Pro	Val	Val.: 210	Ser	Thr (Gln	Leu	Leu 215	Leu .	Asn (Gly	Ser	Leu 220	
45	GCA (Glu (Glu /	Asp	Ile ' 225	Val	Ile	Arg	Ser i	Ala <i>i</i> 230	Asn 1	Leu	Thr	
		asn / 235	Ala I	-ys	Asn :	Ile	Ile 240	Val (Gln 1	Leu A	Asn (31u 245	Ser	
50		rnr)	det A	Asn (250	Cys 1	Chr .	Arg	Pro i	Asn <i>l</i> 255	Asn A	sn 1	Thr :	Met	
55	Lys S 260	ser]	ile E	lis :	Cle C	31y 1 265	Pro	Gly I	Arg A	Ala P	he 1	Tyr i	Ala	
J	ACA C	, 1y <i>F</i>	15n 1 175	le 1	(le G	ly A	Asp	Ile / 280	Arg G	iln A	la F	lis (Cys 285	
60	AAC A	ile S	er G	21y 1	thr L	ys 1	Crp /	Asn A	sp T	hr L 195	eu [ys l	Lys	
	ATA CILE A	la I	TA A	AA 1 ys L	TA A .eu A	rg (GAA (Glu (BOS	CAA 1 Gln F	TT A	AT A	ys T	CA A	ATA 9	33

```
GTC TTT AAT CAA TCC TCA GGA GGG GAC CCA GAA ATT GCA 972
      Val Phe Asn Gln Ser Ser Gly Gly Asp Pro Glu Ile Ala
                                     320
                 315
      ACG CTC AGT TTT AAT TGT GGA GGG GAA TTT TTC TAC TGT 1011
      Thr Leu Ser Phe Asn Cys Gly Gly Glu Phe Phe Tyr Cys
 5
                         330
                                             335
      325
      AAT TCA ACA CAA CTG TTT AAT AGT ACT TGG AAT AGT ACT 1050
      Asn Ser Thr Gln Leu Phe Asn Ser Thr Trp Asn Ser Thr
                                 345
             340
      GGG TCA AAT AAC ACT AAA GGA AAT GAC ACA ATC ACA CTC 1089
10
      Gly Ser Asn Asn Thr Lys Gly Asn Asp Thr Ile Thr Leu
                                         360
                     355
      CCA TGC AGA ATA AGA CAA ATT ATA AAC ATG TGG CAG AAA 1128
      Pro Cys Arg Ile Arg Gln Ile Ile Asn Met Trp Gln Lys
                              370
                                                 375
         365
15
      ATA GGA AAA GCA ATG TAT GCC CCT CCC ATC AAA GGG CAA 1167
      Ile Gly Lys Ala Met Tyr Ala Pro Pro Ile Lys Gly Gln
                 380
                                     385
      ATT AGA TGT TCA TCA AAT ATT ACA GGG CTA ATA TTA ACA 1206
      Ile Arg Cys Ser Ser Asn Ile Thr Gly Leu Ile Leu Thr
20
                         395
                                             400
      390
      AGA GAT GGT GGT AAC AAC AAC ATG AGC AAG ACC ACC GAG 1245
      Arg Asp Gly Gly Asn Asn Asn Met Ser Lys Thr Thr Glu
             405
                                 410
                                                     415
      ACC TTC AGA CCT GGA GGA GGA GAT ATG AGG GAC AAT TGG 1284
25
     Thr Phe Arg Pro Gly Gly Gly Asp Met Arg Asp Asn Trp
                     420
                                         425
     Arg Ser Glu Leu Tyr Lys Tyr Lys Val Val Lys Ile Glu
                                                440
                             435
30
         430
      CCA TTA GGA GTA GCA CCC ACC AGG GCA AAG AGA AGA GTG 1362
      Pro Leu Gly Val Ala Pro Thr Arg Ala Lys Arg Arg Val
                                     450
                 445
     GTG CAG AGA GAA AAA AGA GCA GTG GGA ATA GGA GCT GTG 1401
     Val Gln Arg Glu Lys Arg Ala Val Gly Ile Gly Ala Val
35
                         460
                                             465
     455
      TTC CTT GGG TTC TTG GGA GCA TAA AGC TTC TAG A 1435
     Phe Leu Gly Phe Leu Gly Ala Xaa Ser Phe Xaa
             470
```

40

45

50

In addition to the listing in Table 1, Figure 3 shows the alignment of the amino acid sequences of the clones of each of the seven isolates. Corresponding residues from various clones are in boxes. In the figure, the amino acid sequences are aligned against MN-rgp120 (SEQ. ID. NO. 29).

In one embodiment, a gp120 polypeptide of this invention has the same amino acid sequence as the sequence of one of the breakthrough isolates. In another embodiment, the amino acid sequence is truncated, as described in detail hereinafter. In another embodiment, a gp120 polypeptide sequence of this invention contains a substitution, insertion, or

deletion (alteration) of one or more amino acids in the sequence of a breakthrough isolate. Usually, with the exception of amino acids that are not present in a truncated amino acid sequence and eliminate an epitope, a gp120 polypeptide of this invention will include alterations in the amino acid sequence of a breakthrough isolate that do not alter the polypeptide's ability to induce the same neutralizing antibodies as the amino acid sequence of the isolate.

In general, substitutions in the amino acid sequence of a gp120 polypeptide of this invention are conservative substitutions, particularly for amino acid residues in the V2, V3, and C4 domains of gp120, which domains contain neutralizing epitopes. However, non-conservative substitutions, particularly in domains that do not contain neutralizing epitopes are contemplated.

10

15

20

25

Conservative substitutions replace an amino acid with an amino acid of similar size and character. For example, a hydrophobic residue or hydrophilic residue is replaced with another hydrophobic residue or hydrophilic residue, respectively. Amino acids can be divided into the following groups: positively charged residues (K, R and H); negatively charged residues (D and E); amides (N and Q); aromatics (F, Y, and W); hydrophobics (P, G, A, V, L, I, and M); and uncharged residues (S and T). Usually, residues within a group are replaced with another member of the group.

In one embodiment, critical amino acid residues in the V2, V3, and C4 domains of gp120 are identical to the corresponding residues in a breakthrough isolate sequence. Critical amino acid residues in the V2, V3, and C4 domains of gp120 are described in the experimental section. In another embodiment, all amino

acid residues in the V2, V3, and C4 domains of gp120 are identical to corresponding residues in a breakthrough isolate sequence.

5 Oligonucleotide Encoding gp120 from Breakthrough Isolates

10

15

20

25

30

35

The present invention also provides novel oligonucleotides encoding gp120 from the breakthrough isolates which can be used to express gp120. An oligonucleotide of this invention encodes a polypeptide of this invention. The oligonucleotide can be DNA or RNA, usually DNA. Although numerous nucleotide sequences can encode the same amino acid sequence due to the degeneracy of the genetic code, conveniently, the oligonucleotides of this invention include a nucleotide sequence of a breakthrough isolate as illustrated in Table 1 (Sequence ID Nos. 2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, and 28). Usually, an oligonucleotide of this invention is less than about 5 kilobases (kb), preferably less than about 3 kb.

To express the encoded amino acid sequence, the oligonucleotide can be inserted into a transcription The transcription unit can be inserted into a plasmid for production of cell lines, inserted into a virus (e.g.; vaccinia) or can be used directly as a DNA vaccine. Suitable transcription units for production of vaccine proteins are well known. A preferred expression vector, designated psvI6B5, is illustrated in Sequence ID No. 32. The vector includes an HSV-1 gD1 signal sequence joined to a linker sequence. The gp120 nucleotide sequence to be expressed starts with the Kpn I site of the gene. Since all gp120 or gp160 sequences contain this site, any gp120 nucleotide sequence can be analogously inserted into the vector and expressed. The vector ends with a polyA tail from SV40.

In addition to being useful to express a polypeptide sequence of this invention, the oligonucleotides of this invention can also be used in diagnostics to detect HIV isolates. For example, the oligonucleotide or a portion thereof encoding a neutralizing epitope can be used in branched chain DNA diagnostics or as a probe in in situ hybridization studies.

10 <u>Vaccine preparation</u>

5

25

30

35

A gp120 polypeptide of this invention from a selected breakthrough isolate(s) in a suitable carrier is used to make a subunit vaccine. The polypeptide can be used alone, but is generally administered in a multivalent subunit vaccine that includes gp120 MN. In addition to one or more gp120 polypeptides of this invention, the vaccine generally includes the MN polypeptide (hereinafter, MN-rgp120). The vaccine usually includes about 3 to about 5 different gp120 polypeptides, but 30 or more different gp120 polypeptides can be used.

Preparation of gp120 polypeptides for use in a vaccine is well known and is described hereinafter. With the exception of the use of the selected HIV isolate, the gp120 subunit vaccine prepared in the method does not differ from gp120 subunit vaccines of the prior art.

As with prior art gp120 subunit vaccines, gp120 at the desired degree of purity and at a sufficient concentration to induce antibody formation is mixed with a physiologically acceptable carrier. A physiologically acceptable carrier is nontoxic to a recipient at the dosage and concentration employed in the vaccine. Generally, the vaccine is formulated for injection, usually intramuscular or subcutaneous injection. Suitable carriers for injection include

sterile water, but preferably are physiologic salt solutions, such as normal saline or buffered salt solutions such as phosphate-buffered saline or ringer's lactate. The vaccine generally contains an adjuvant. Useful adjuvants include QS21 (Quillaja saponaria, commercially available from Cambridge Biotech, Worcester, MA), which stimulates cytotoxic T-cells, and alum (aluminum hydroxide adjuvant). Formulations with different adjuvants which enhance cellular or local immunity can also be used. In particular, immunopotentiators such as cytokines can be included in the vaccine. Examples of suitable immunopotentiating cytokines include interleukins, such as interleukin-2 (IL-2) and interleukin-12 (IL-12), and tumor necrosis factor-alpha (TNF-α).

10

15

20

25

30

35

Additional excipients that can be present in the vaccine include low molecular weight polypeptides (less than about 10 residues), proteins, amino acids, carbohydrates including glucose or dextrans, chelating agents such as EDTA, and other excipients that stabilize the protein or inhibit growth of microorganisms.

The vaccine can also contain other HIV proteins. In particular, gp41 or the extracellular portion of gp41 or HIV-1 core proteins such as P24, P17, and P55 can be present in the vaccine. Although the amino acid sequence of gp41 is more conserved than that of gp120, gp41 contains neutralizing epitopes. Preferably, any gp41 present in the vaccine is from an HIV isolate present in the vaccine. gp160 from an isolate used in the vaccine can replace gp120 in the vaccine or be used together with gp120 from the isolate. Alternatively, gp160 from a different isolate than those in the vaccine can additionally be present in the vaccine.

Vaccines according to the invention can also contain one or more soluble gp120 polypeptide

sequences, or fragments thereof, in combination with an engineered virus specifically designed to express proteins that induce a cytotoxic T-cell response. Suitable engineered viruses are derived from, for example, Canary Pox virus, vaccinia viruses, attenuated human herpes viruses (such as, e.g., herpes simplex viruses), and Varicella Zoster. Exemplary engineered viruses are modified to express any HIV protein capable of inducing a cytotoxic T-cell response, such as those described above. Typically, immunization with the gp120/engineered virus vaccine is followed by administration of one or more doses of the gp120 polypeptide sequence(s) to boost the immune response. If desired, viruses can be engineered to express one or more gp120 polypeptide sequences of the invention, or fragments thereof, and used in vaccines with or without soluble gp120 polypeptide sequences.

10

15

20

25

30

35

Vaccine formulations generally include a total of about 300 to 600 μ g of gp120, conveniently in about 1.0 ml of carrier. Preferred formulations include use of twice the weight of a gp120 polypeptide in twice as600 μ g alum. However, formulations having smaller amounts (e.g.; 50 μ g per dose) are also used, generally with alum or other adjuvants. The amount of gp120 for any isolate present in the vaccine will vary depending on the immunogenicity of the gp120. For example, gp120 from some strains of HIV may be less immunogenic than gp120 from the MN strain (Sequence ID No. 29). strains having different immunogenicity are used in combination, empirical titration of the amount of each virus would be performed to determine the percent of the gp120 of each strain in the vaccine. For isolates having similar immunogenicity, approximately equal amounts of each isolate's gp120 would be present in the vaccine. For example, in a preferred embodiment, the vaccine includes gp120 from the MN and a strain of this

invention at concentrations of about 300 μ g per strain in about 1.0 ml of carrier. When the vaccine includes gp120 from about 30 isolates, about 10 to about 50 μ g can be used. Methods of determining the relative amount of an immunogenic protein in multivalent vaccines are well known and have been used, for example, to determine relative proportions of various isolates in multivalent polio vaccines.

The vaccines of this invention are administered in the same manner as prior art HIV gp120 subunit vaccines. In particular, the vaccines are generally administered at 0, 1, and at 6, 3 or 12 months, depending on the protocol. A preferred protocol includes administration at 0, 1, 6, and 12 months.

Following the immunization procedure, annual or bi-annual boosts can be administered. However, during the immunization process and thereafter, neutralizing antibody levels can be assayed and the protocol adjusted accordingly.

The vaccine is administered to uninfected individuals. In addition, the vaccine can be administered to seropositive individuals to augment immune response to the virus, as with prior art HIV vaccines. It is also contemplated that DNA encoding the strains of gp120 for the vaccine can be administered in a suitable vehicle for expression in the host. In this way, gp120 can be produced in the infected host, eliminating the need for repeated immunizations. Preparation of gp120 expression vehicles is described hereinafter.

20

25

30

35

Although the gp120 isolates described herein can be used as a vaccine as described above, the amino acid sequences can also be used alone or in combinations in the same type of formulation for use as an immunogen, to induce antibodies that recognize the isolate(s) present in the immunogen. Immunogens are formulated in

the same manner as vaccines and can include the same excipients, etc. Antibodies induced by the immunogens can be used in a diagnostic to detect the HIV strain in the immunogen or to affinity purify the strain.

5

10

15

20

25

30

35

gp120 Polypeptide Sequences and Chemokine Receptors

While CD4 is the primary cellular receptor for HIV-1, it is not sufficient for entry of HIV-1 into cells. Co-receptors required in conjunction with CD4 have been identified. These co-receptors are members of the chemokine receptor family of seven-transmembrane G-protein coupled receptors. The chemokine superfamily is subdivided into two groups based on the amino terminal cysteine spacing. The CXC chemokines are primarily involved in neutrophil-mediated inflammation, and the CC chemokines tend to be involved in chronic inflammation. At least five CC chemokine receptors, designated CC-CKR1-5 (also known in the art as CCR1-5), and at least four CXC chemokine receptors, designated CXC-CKR1-4 (also known as CXCR-1-4), have been identified.

CXC-CKR-4 (CXCR-4), which has also been called the alpha-chemokine receptor fusin, serves as an entry cofactor for T-cell-tropic HIV-1 strains. CC-CKR-5 (CC-R5), which has been called beta-chemokine receptor, together with its related family members, such as CC-CKR-2b and CC-CKR3, serve as entry cofactors for macrophage-tropic HIV-1 strains. T-cell-tropic strains can infect primary T-cells and T-cell lines, but not macrophages, whereas macrophage-tropic strains can infect macrophages and primary T-cells, but not T-cell lines. T-cell- and macrophage-tropic strains are discussed more fully in Deng et. al., Nature 381:661-666 (1996), which is hereby incorporated by reference in its entirety. Examples of T-cell-tropic strains include laboratory isolates, such as IIIB and MN.

Macrophage-tropic strains include primary isolates, including but not limited to A244, GNE6, GNE8, and breakthrough viruses from vaccinees immunized with gp120-based vaccines. Dual-tropic strains can use both types of co-receptors, entering cells via CXC-CKR-4 or via one or more CC-CKR family members, preferably CC-CKR-5, CC-CKR-2b, or CC-CKR-3. While the present invention is not intended to be bound or limited by any one theory, the entry of T-cell tropic and macrophage-tropic HIV-1 strains is believed to provide a unifying explanation of the differences in cell tropism between viral strains, the resistance to HIV-1 infection by many CD4-transfected nonprimate cells, and the HIV-1-infection resistance of a portion of the human population.

5

10

15

20

25

30

35

Accordingly, in one embodiment is a vaccine containing (1) a first gp120 polypeptide sequence, or fragment thereof, from a macrophage-tropic HIV-1 strain and/or a second gp120 polypeptide sequence, or fragment thereof, from a T-cell tropic strain, in combination with (2) a breakthrough isolate HIV gp120 polypeptide sequence, or fragment thereof, from a vaccinee vaccinated with the first and/or second HIV gp120 polypeptide sequence. Preferably, the vaccine includes at least two gp120 polypeptide sequences that bind to different chemokine receptors. In one embodiment, the vaccine includes first and second gp120 polypeptide sequences that bind to different chemokine receptors. In addition, the breakthrough isolate gp120 polypeptide sequence can bind to a different chemokine receptor than the chemokine receptor(s) bound by either or both of the first and second gp120 polypeptide sequence(s).

A preferred T-cell tropic strain is a laboratory isolate, most preferably MN. Preferred macrophage-tropic viruses for use in the invention are GNE6 and GNE8, which are representative of the breakthrough

viruses disclosed herein and differ from MN in that their gpl20s induce the formation of antibodies that recognize the gpl20 sequences (e.g., the V3 domain) involved in binding to CC chemokine receptors, such as CXC-CKR-5.

5

10

15

25

30

35

In one embodiment, HIV infection is prevented by administering one or more chemokine receptor-binding gp120 polypeptide sequences, or fragment(s) thereof containing appropriate chemokine receptor-binding domains, in a vaccine, such as those described above. Preferably, the vaccine also includes one or more CD4-binding gp120 polypeptide sequences or appropriate fragments thereof. Such vaccines induce anti-HIV antibodies that inhibit viral gp120-chemokine receptor or -CD4 binding. In addition, such gp120 polypeptides can directly inhibit HIV infection by binding to one or more co-receptors for HIV infection, such as CD4 or a chemokine receptor, thus providing a prophylactic or therapeutic effect in treating HIV infection.

Preferably, gp120 polypeptide sequences useful in this regard contain the T-cell binding (TCB) domain.

Various uses of chemokine receptor-binding gp120 polypeptides are discussed below with regard to the CC chemokine receptor family. However, those skilled in the art recognize that this discussion applies equally to CXC chemokine receptors that act as cofactors in HIV infection.

The gp120 polypeptides can be used as a composition containing one or more gp120 polypeptides, as described for use as a vaccine or immunogen. The composition can be administered, prophylactically or therapeutically, to a patient at risk of infection or in need of such treatment using the dosages and routes and means of administration described herein. However, chronic administration may be preferred and dosages can be adjusted accordingly. It is noted that in vivo

administration can also induce antibodies that bind viral gp120, further inhibiting virus binding to CC-CKR.

5

10

15

The gp120 polypeptides can also be used in screening assays to identify antagonists of CC-CKR. For example, candidate antagonists can be screened for inhibition of binding of gp120 to a CC-CKR CC-CKR receptor that is isolated and attached to a surface (e.g., plastic dish) or recombinantly or naturally expressed on the surface of a cell. Antagonists can either bind gp120 or bind receptor. Preferred candidate antagonists include gp120 compounds, small gp120 peptides (5 to 20 amino acids in length, preferably 7 to 10 amino acids in length) or peptidomimetics of gp120 that bind receptor, monoclonal antibodies that bind gp120, and small organic molecules that bind either gp120 or receptor.

The antibodies induced by the gp120 polypeptides can also be used to induce anti-idiotype antibodies 20 that bind CC chemokines. These anti-idiotype antibodies can be screened for binding to an anti-gp120 polypeptide antibody and inhibiting gp120 from binding CC-CKR receptor. Such anti-idiotype antibodies mimic gp120 by binding to CC-CKR receptor. Such antibodies, preferably human antibodies, can be obtained in a 25 number of ways, such as human antibodies from combinatorial libraries (e.g., Burton et al. Adv. Immunolo. (1994) 57:191-280). It is now possible to produce transgenic animals (e.g., mice) that are capable, upon immunization, of producing a full 30 repertoire of human antibodies in the absence of endogenous immunoglobulin production. For example, homozygous deletion of the antibody heavy-chain joining region (JH) gene in chimeric and germ-line mutant mice results in complete inhibition of endogenous antibody 35 production. Transfer of the human germ-line

immunoglobulin gene array in such germ-line mutant mice results in the production of human antibodies upon antigen challenge as described in Jakobovitis et al., Proc. Natl. Acad. Sci. USA 90: 2551 (1993); Jakobovits et al., Nature 362:255-258 (1993); Bruggermann et al., Year in Immuno. 7: 33 (1993).

Alternatively, phage display technology as described by McCafferty et al., Nature 348:552-553 (1990) can be used to produce human antibodies and antibody fragments in vitro from immunoglobulin variable (V) domain gene repertoires from unimmunized donors. According to this technique, antibody V domain genes are closed in-frame either into either a major or minor coat protein gene of a filamentous bacteriophage, such as M13 or fd, and displayed as functional antibody fragments on the surface of the phage particle. Because the filamentous particle contains a singlestranded DNA copy of the phage genome, selections based on the functional properties of the antibody also result in selection of the gene encoding the antibody exhibiting those properties. Phage display can be performed in a variety of formats as reviewed by, for example, Johnson, et al., Current Opinion in Structural Biology 3:564-571 (1993).

10

15

20

Several sources of V-gene segments can be used for phage display. Clackson et al., Nature, 352: 624-628 (1991) isolated a diverse array of anti-oxazolone antibodies from a small random combinatorial library of V genes derived from the spleens of immunized mice. A repertoire of V genes from unimmunized human donors (or embryonic cells) can be constructed. It has been demonstrated that antibodies to a diverse array of antigens (including self-antigens) can be isolated essentially following the techniques described by Marks et al., J. Mol. Biol., 222: 581-597 (1991), or Griffith et al., EMBO J., 12: 725-734 (1993).

In a natural immune response, antibody genes accumulate mutations at a high rate (somatic hypermutation). Some of the changes introduced confer higher affinity, and B cells displaying high-affinity surface immunoglobulin are preferentially replicated and differentiated during subsequent antigen challenge. This natural process can be mimicked by employing the technique known as "chain shuffling" (Marks et al., Bio/Technol. 10:779-783 [1992]). In this method, the affinity of "primary" human antibodies obtained by phage display can be improved by sequentially replacing the heavy and light chain V region genes with repertoires of naturally occurring variants (repertoires) of V domain genes obtained from unimmunized donors. This technique allows the production of antibodies and antibody fragments with affinities in the nM range. A strategy for making very large phage antibody repertoires has been described by Waterhouse et al., Nucl. Acids Res., 21: 2265-2266 (1993).

10

15

20

25

30

35

Accordingly, antibodies that bind CC-CKR can be obtained by screening antibodies or fragments thereof expressed on the surface of bacteriophage in combinatorial lfbraries or in other systems as described above with a gp120 monoclonal antibody that inhibits gp120 binding to receptor.

In addition to screening antibodies with a gp-120 antibody, random or combinatorial peptide libraries can be screened with either a gp120 antibody or the gp120 compounds of the invention. Approaches are available for identifying peptide ligands from libraries that comprise large collections of peptides, ranging from 1 million to 1 billion difference sequences, which can be screened using monoclonal antibodies or target molecules. The power of this technology stems from the chemical diversity of the amino acids coupled with the

large number of sequences in a library. See for example, Scott et al., Cur. Open. Biotechnol. 5(1):40-8 (1994); Kenan et al. Trends Biochem. Sci. (1994) 19(2):57-64. Accordingly, the monoclonal antibodies, preferably human monoclonals or fragments thereof, generated as discussed herein, find use in treatment by inhibiting or treating HIV infection or disease progression, as well as in screening assays to identify additional pharmaceuticals.

10

15

20

30

35

Production of gp120

gp120 for a vaccine can be produced by any suitable means, as with prior art HIV gp120 subunit vaccines. Recombinantly-produced or chemically synthesized gp120 is preferable to gp120 isolated directly from HIV for safety reasons. Methods for recombinant production of gp120 are described below.

Oligonucleotides encoding gp120 from breakthrough isolates and capable of expressing gp120 can be prepared by conventional means. For example, the nucleotide sequence can be synthesized. Alternatively, another HIV nucleotide sequence encoding gp120 can be used as a backbone and altered at any differing residues as by site-directed mutagenesis.

25 Site-directed mutagenesis is described in Kunkel et al, Proc. Natl. Acad. Sci. (USA) 82:488-492 (1985) and Zoller et al, Nuc. Acids Res. 10:6487-6500 (1982) and is well known.

In a preferred embodiment, the nucleotide sequence is present in an expression construct containing DNA encoding gp120 under the transcriptional and translational control of a promoter for expression of the encoded protein. The promoter can be a eukaryotic promoter for expression in a mammalian cell. In cases where one wishes to expand the promoter or produce gp120 in a prokaryotic host, the promoter can be a

prokaryotic promoter. Usually a strong promoter is employed to provide high-level transcription and expression.

5

10

15

20

25

30

35

The expression construct can be part of a vector capable of stable extrachromosomal maintenance in an appropriate cellular host or may be integrated into host genomes. Normally, markers are provided with the expression construct which allow for selection of a host containing the construct. The marker can be on the same or a different DNA molecule, desirably, the same DNA molecule.

The expression construct can be joined to a replication system recognized by the intended host cell. Various replication systems include viral replication systems such as those from retroviruses, simian virus, bovine papilloma virus, or the like. In addition, the construct may be joined to an amplifiable gene, e.g. the DHFR gene, so that multiple copies of the gp120 DNA can be made. Introduction of the construct into the host will vary depending on the construct and can be achieved by any convenient means. A wide variety of prokaryotic and eukaryotic hosts can be employed for expression of the proteins.

preferably, the gp120 is expressed in mammalian cells that provide the same glycosylation and disulfide bonds as in native gp120. Expression of gp120 and fragments of gp120 in mammalian cells as fusion proteins incorporating N-terminal sequences of Herpes Simplex Virus Type 1 (HSV-1) glycoprotein D (gD-1) is described in Lasky, L. A. et al., 1986 (Neutralization of the AIDS retrovirus by antibodies to a recombinant envelope glycoprotein) Science 233: 209-212 and Haffar, O.K. et al., 1991 (The cytoplasmic tail of HIV-1 gp160 contains regions that associate with cellular membranes.) Virol. 180:439-441, respectively. A preferred method for expressing gp120 is described in

the examples. In the examples, a heterologous signal sequence was used for convenient expression of the protein. However, the protein can also be expressed using the native signal sequence.

An isolated, purified gp120 polypeptide having one of the amino acid sequences illustrated in Table 1 can be produced by conventional methods. For example, the proteins can be chemically synthesized. In a preferred embodiment, the proteins are expressed in mammalian cells using an expression construct of this invention. The expressed proteins can be purified by conventional means. A preferred purification procedure is described in the examples.

15 <u>gp120 Fragments</u>

5

10

20

25

30

35

The present invention also provides gp120 fragments that are suitable for use in inducing antibodies for use in a vaccine formulation. A truncated gp120 sequence, as used herein, is a fragment of gp120 that is free from a portion of the intact gp120 sequence beginning at either the amino or carboxy terminus of gp120. A truncated gp120 sequence of this invention is free from the C5 domain. The C5 domain of gp120 is a major immunogenic site of the molecule. However, antibodies to the region do not neutralize virus. Therefore, elimination of this portion of gp120 from immunogens used to induce antibodies for serotyping is advantageous.

In another embodiment, the truncated gp120 sequence is additionally free from the carboxy terminal region through about amino acid residue 453 of the gp120 V5 domain. The portion of the V5 domain remaining in the sequence provides a convenient restriction site for preparation of expression constructs. However, a truncated gp120 sequence that is free from the entire gp120 V5 domain is also

suitable for use in inducing antibodies.

In addition, portions of the amino terminus of gp120 can also be eliminated from the truncated gp120 In particular, the truncated gp120 sequence sequence. can be free from the gp120 signal sequence. truncated gp120 sequence can be free from the carboxy terminus through amino acid residue 111 of the gp120 C1 domain, eliminating most of the Cl domain but preserving a convenient restriction site. However, the portion of the C1 domain through the V2 cysteine residue that forms a disulfide bond can additionally be removed, so that the truncated gp120 sequence is free from the carboxy terminus through amino acid residue 117 of the gp120 C1 domain. In a preferred embodiment, the truncated gp120 sequence is free from the amino terminus of gp120 through residue 111 of the C1 domain and residue 453 through the carboxy terminus of gp120.

The truncated gp120 sequences can be produced by recombinant engineering, as described previously. Conveniently, DNA encoding the truncated gp120 sequence is joined to a heterologous DNA sequence encoding a signal sequence.

It is understood that the application of the teachings of the present invention to a specific problem or situation is within the capabilities of one having ordinary skill in the art in light of the teachings contained herein. Examples of the products of the present invention and representative processes for their isolation, use, and manufacture appear below, but should not be construed to limit the invention. All literature citations herein are expressly incorporated by reference.

35

5

10

15

20

25

30

EXAMPLES
Materials and Methods

Specimen collecti n from human volunt ers. was collected from MN-rgp120-immunized individuals who were infected with HIV-1 while participating in Phase I (NIH Protocol AVEG 016) and Phase II (NIH Protocol AVEG 201) HIV-1 vaccine trials sponsored by the 5 National Institutes of Health (NIH). The demographics of the subjects in the study, and the study design have been described in McElrath; Seminars in Cancer Biol. 6:1-11 (1995); McElrath et al.; Abstracts from Eighth Annual Meeting of the National Cooperative Vaccine 10 Development Groups for AIDS. Bethseda, MD 216 (1996). Specimens were obtained according to an informed consent protocol approved by the institutional review boards of the participating institutions. In the experimental section, the time of HIV-1 infection is 15 specified with regard to data provided by the NIH AIDS Vaccine Evaluation Network where PCR (RNA) and/or serologic assays were used to detect HIV-1 infection.

Sample preparation for cloning HIV-1 envelope 20 glycoproteins. Peripheral blood mononuclear cells (PBMCs) from HIV-1 infected vaccinees were prepared from heparinized venous blood by FICOLL-HYPAQUE gradient centrifugation. Cell number and viability were determined. After separation, PBMCs were washed 25 twice in phosphate-buffered saline and suspended at a cell density of 6x10° cells/ml in PCR lysis buffer (50 mM KCl, 10 mM Tris (pH 8.4), 2.5 mM MgCl2, 0.1 mg/ml gelatin (Sigma), 0.45% NONIDET P40 detergent, 0.45% TWEEN 20 detergent (both detergents are 30 commercially available from United States Biochemical Corp.) and 0.06 mg/ml Proteinase K (Gibco BRL) to lyse the cells. The lysate was incubated at 50-60°C for 1 hour, followed by inactivation of the Proteinase K at 95°C for 10 minutes. Lysates were shipped frozen and 35 stored at -70°C until use.

Polymerase chain reaction (PCR) amplificati n.

Samples were subjected to two rounds of PCR
amplification using the nested primers described below.

In the first round, 25 µl aliquots of PBMC lysates

(containing about 1 µg genomic DNA) were mixed with an equal volume of a PCR reaction mix containing 400 µM
each dNTP, 200 µg/ml BSA (Sigma Chemical Corporation,
RIA grade) and about 100 pmoles of each primer in 50 mM
KCl, 20 mM Tris (pH 8.4) and 3 mM MgCl₂. After an
initial 10 minute denaturation step at 95°C, 5 units of
Taq polymerase (AMPLITAQ, Perkin Elmer Cetus) were
added during an 55°C soak step, and samples were
overlayed with mineral oil.

The PCR profile was as follows: 2 cycles having 1 minute at 55°C, 2.5 minutes at 72°C and 1 minute at 94°C, followed by 28 cycles with 30 seconds at 55°C, 2.5 minutes at 72°C and 45 seconds at 94°C, and an extension step at 72°C for 5 minutes.

15

20

25

Aliquots of 10 µl from the first-round reactions were re-amplified with appropriate nested primers in a final reaction volume of 100 µl, using either the reagents and profile described above or the reagents and profile described in the PCR Optimizer Kit (Invitrogen.) PCR reaction products were purified using QIAQUICK-spin columns (Qiagen Inc.) The primer pair used in the first round was either 120.os.F (5'-gggaattcggatccAGAGCAGAAGACAGTGGCAATGA with homologous sequence at position 6248-6270 of HIVPV22) (SEQ. ID. NO. 34) or JM11A

30 (5'-ctcgag-CTCCTGAAGACAGTCAGACTCATCAAG at position
6048-6074) (SEQ. ID. NO. 35) in the forward direction
[Kusumi et al.; J. Virol. 66:875 (1992)] combined with
120.os.R (5'-ggtctagaagctttaGCCCATAGTGCTTCCTGCTGCT-CC
at position 7836-7859) (SEQ. ID. NO. 36) in the reverse
35 direction. The internal nested primers were 120.BX.F
(5'-gggcggatcctcgaGGTACCTGTRTGGAAAGAAGCA at position

6389-6410; R: A or G) (SEQ. ID. NO. 37) and 120.is.R (5'-ggtctagaagctttaTGCTCCYAAGAACCCAAGGAACA at position 7819-7841; Y: T or C) (SEQ. ID. NO. 38). Heterologous primer sequences are shown in lower case letters.

5

10

15

20

25

35

Subcloning of PCR products and the expression of recombinant envelope glycoproteins as fusion proteins. The HIV-1 envelope glycoprotein gp120 sequences were cloned and expressed as chimeric genes and fusion proteins, where the signal sequence and 27 amino acids from the mature N terminus of herpes simplex virus type 1 (HSV-1) were fused to the N-terminal sequences of the gp120 genes, corresponding to amino acid 13 of the mature gp120 sequence. PCR products containing gp120 sequences from the breakthrough specimens were cloned into pRK5 expression plasmid as chimeric genes using combinations of restrictions sites engineered into the heterologous PCR primer tails and the Xho I site engineered into the N-terminal sequence of HSV-1 qD.

The resulting double-stranded DNA was sequenced with Sequenase and the dGTP Reagent Kit (United States Biochemical Corp.). Sequences from glycoprotein D were provided to enhance expression and to provide a flag epitope to facilitate protein analysis, as described in Berman et al.; J. Virol. 7:4464-9 (1992); Nakamura et al.; AIDS and Human Retroviruses 8:1875-85 (1992); and Nakamura et al.; J. Virol. 67:6179-91 (1993).

Briefly, isolated DNA fragments generated by the PCR reaction were ligated into a plasmid (pRK.gD-5, 30 pRKgDstop) designed to fuse the gp120 fragments, in frame, to the 5' sequences of the glycoprotein D (gD) gene of Type 1 Herpes Simplex Virus (gD-1) and the 3' end to translational stop codons. The fragment of the gD-1 gene encoded the signal sequence and 25 amino acids of the mature form of HSV-1 protein. To allow

PCT/US97/09690 WO 98/01564

for expression in mammalian cells, chimeric genes fragments were cloned into the pRK5 expression plasmid (Eaton et al., Biochemistry 291:8343-8347 (1986)) that contained a polylinker with cloning sites and translational stop codons located between a cytomegalovirus promotor and a simian virus 40 virus polyadenylation site.

The resulting plasmids were transfected into the 293s embryonic human kidney cell line (Graham et al., J. Gen. Virol. 36:59-77 (1977)) using a calcium 10 phosphate technique (Graham et al., Virology 52:456-467 (1973)). Growth conditioned cell culture media was collected 48 hr after transfection, and the soluble proteins were detected by ELISA or by specific radioimmunoprecipitation where metabolically labeled 15 proteins from cell culture supernatants were resolved by sodium dodecyl sulfate polyacrylamide gel electrophoresis (PAGE) and visualized by autoradiography as described in Berman et al., J. Virol. 63:3489-3498 (1989) and Laemmli, Nature 20 227:680-685 (1970).

Serologic assays. Sera were assayed for antibodies to rgp120, antibodies to synthetic gp120 V3 domain peptides corresponding to sequences from the qp120 V3 domain, and antibodies able to inhibit the binding of MN-rgp120 to cell surface CD4 using serologic assays described in Berman et al.; J. Virol. 7:4464-9 (1992); Nakamura et al.; AIDS and Human Retroviruses 8:1875-85 (1992); and Nakamura et 30 al.; J. Virol. 67:6179-91 (1993). Endpoint titers of antibody binding to gp120 and V3 peptides were determined using three fold-serial dilutions of sera. The endpoint dilution titer was defined as the last dilution that produced an optical density value that was two times higher than the mean of the optical

25

35

densities of 1:50 diluted, pooled, normal human sera. Antibody titers were calculated by a computer program that interpolated values between antibody dilutions. The inter-assay coefficient of variation of positive control standard sera was 35%.

Binding of monoclonal antibodies to rgp120 from breakthrough viruses. An ELISA similar to that described by Moore et al.; AIDS 3:155-63 (1989) was used to measure the binding of various monoclonal 10 antibodies (MAbs) to rgp120s from breakthrough viruses. Briefly, Nunc-Immuno plates (Maxisorp, certified) were coated (100 μ l at 5 μ g/ml in PBS at 4°C overnight) with an affinity-purified sheep polyclonal antiserum to a peptide at the C terminus of gp120 (D7324, 15 International Enzymes, Fallbrook, CA). After washing once with PBS-0.05% TWEEN-20 detergent, the plates were blocked with PBS-1.0% BSA for 30-60 minutes at room temperature. Cell culture supernatants from 293s cells, diluted to contain equivalent amounts of the 20 gD-rgp120 fusion protein, were added and incubated for 2 hours at room temperature followed by three washes with PBS-0.05% TWEEN-20 detergent. Various MAbs were diluted in PBS-1.0% BSA and 100 μL of the diluted MAbs 25 were added to each well and incubated for 1 hour at room temperature.

The plates were washed 3 times and incubated with 100 μ l of a horseradish peroxidase-conjugated second antibody (goat anti-mouse or anti-human IgG, Cappel) for 1 hour at room temperature. After 3 washes the plates were developed and the OD₁₀, (optical density at 492 nm) read in a plate reader. Growth conditioned cell culture supernatants were normalized by dilution based on binding by MAb 5B6 which is specific for HSV-1 glycoprotein D fusion protein.

30

35

Virus n utralization assays. The ability of vaccinee sera to inhibit infection of MT4 cells by HIV-1_{MN} was measured in a cytopathicity assay where cell viability was quantitated using a calorimetric indicator dye, as described in Robertson et al.; J. Virol. Methods 20:195-202 (1988). Briefly, a virus stock of HIV-1_{MN} (obtained from Dr. Michael Norcross, U.S. Food and Drug Administration) was prepared as the clarified supernatant from chronically infected H9/HIV-1_{MN} cell culture. H9 cells chronically infected with HIV-MN were pelleted and resuspended in one-tenth the original volume of medium. Cell-associated virus was released by the mechanical shearing effects of rapid vortexing of the cells as described in Wrin et al.; J. Virol. 69:39-48 (1995).

5

10

15

35

An amount of virus sufficient to ensure complete cell lysis killing in 7 days was incubated with three-fold serial dilutions of test antisera, and then used to challenge MT4 T-lymphoid cells in 10% FCS/RPMI-1640 cell culture media. The cultures were 20 incubated for 7 days at 37°C in 5% CO,, and then cell viability was tested by the dye MTT, as described by Robertson et al.; J. Virol. Methods 20:195-202 (1988). Virus neutralization endpoints were quantitated by measurement of OD at 570-650 nm, and then the endpoint 25 titers were calculated as the reciprocal of the antiserum dilution giving a signal that was two-fold above the control signal with unprotected (killed) cells. These titers were typically twice those calculated at 50% protection. 30

Results

Immunization history of infected subjects. Since 1992, 499 adults have been immunized with MN-rgp120 in Phase I trials in low or moderate risk individuals and in a Phase II clinical trial involving moderate to high

risk individuals. The studies described herein entail the genetic and immunologic characterization of the first seven of nine individuals who became infected with HIV-1 through high risk behavior during the course of these trials. A listing of the trials and summary of the status of the vaccinees is presented in Table 2A. A listing of the analysis of the vaccinees is presented in Table 2B.

10 TABLE 2A

Description of Vaccinees Infected with HIV-1

After Immunization with MN-rgp120

				‡Antigen dose/
	Study No.	Case No.	*Risk Group	<u>Adjuvant</u>
15	016	C6	M/H	300/QS21
	016	C8	M/H	600/QS21
	016	C15	M/H	300/QS21
	201	C 7	M/H	600/Alum
	201	C11	M/H	600/Alum
20	201	C10	M/IDU	600/Alum
	201	C17	M/IDU	600/Alum

^{* -} M/H indicates male homosexual; M/IDU indicate male intravenous drug user.

^{‡ -} numbers indicate dose in micrograms of MN-rgp120 injected per immunization; QS21 indicates antigen was formulated in QS21 adjuvant; Alum indicates MN-rgp120 formulated in aluminum hydroxide.

TABLE 2B

Description of Vaccin s Infected with HIV-1

After Immunization with MN-rgp120

-	Case	Injection Schedule	Injections before	Time of HIV-1+	<pre>pInterval: to HIV-1+</pre>	
5	No.	(months)	HIV-1+	(months)	(months)	
	C 6	0,1,10.5	2	4.00	2.00	
	C8	0,1	2	4.00	3.00	
	C15	0,1,2	3	6.25	4.00	
10	C7	0,1,6,12	3	9.25	3.00	
	C11	0,1,6,12	4	19.50	6.75	
	C10	0,1,6,19	3	19.50	13.50	
	C17	0,1,6,18	4	24.75	6.25	

a - indicates interval between last immunization
 and detection of HIV-1 infection.

Three of the infections occurred in a Phase I trial (NIH Protocol AVEG 201) that compared the safety and immunogenicity of MN-rgp120 formulated in two different adjuvants (alum and QS21), and four of the infections occurred in a Phase II trial aimed at establishing the safety and immunogenicity of MN-rgp120 in various high risk groups (e.g., intravenous drug users, homosexual and bisexual males, and partners of HIV-1 infected individuals).

20.

25

30

Of the seven infections studied (Table 3), two (C6 and C8) occurred after two injections, three (C7, C10 and C15) occurred after three injections, and two (C11 and C17) occurred after receiving the four scheduled injections. The interval between receiving the last immunization and becoming infected was 2 to 13.5 months.

TABLE 3

P ak Post Boost MN-rgp120 Antibody Titers
in Vaccinees that Became Infected with HIV-1

5	Injections	<u>c6</u>	<u>C8</u>	<u>C15</u>	<u>C7</u>	<u>C11</u>	<u>C10</u>	<u>C17</u>
	1	<50	2185	79	<50	1890	na	na
	2	21539	10125	na	413	32696	7771	7056
	. 3	#	,	4460	9707	34728	11627	1841 3
	4	#	#	#	#	#	#	1134
								0

- indicates specimen not analyzed because of
HIV-1 infection.

 $\ensuremath{\text{na}}$ - indicates the sample was not available for testing.

boldface - indicates unusually low antibody
titers.

15

Antibody response to gp120 in vaccinated individuals. The magnitude and specificity of the antibody response to MN-rgp120 was measured by ELISA in all infected individuals throughout the course of the immunization regime (Figure 1). Five of the seven

subjects exhibited normal antibody response kinetics that included a small but reproducible primary response (1:100-1:2,000) and a strong secondary (booster) response (titters ranging from 1:7,000-1:32,000), and antibody responses following third and fourth injections that were similar or marginally higher than those achieved after the second immunization (Figure 1, Table 3).

5

10

15

35

The antibody response observed in C7 (Figure 1C) was unusual in that no antibodies were detectable after the primary injection and a titer of only 1:350 was detected after the second injection. It thus appeared that C7 did not respond to the primary immunization, and that the antibody response obtained after the second injection represented a primary immune response. Consistent with this hypothesis, the third injection elicited a titer of only 1:9,707, typical of those normally seen after two immunizations.

An atypical antibody response was also seen in subject C15 (Figure 1G) who was immunized according to 20 an accelerated immunization schedule of 0, 1, and 2 months. As expected, the antibody titer seen in this subject (1:4,460) was at the low end of what is typically achieved after two immunizations and was far below normal values for three immunizations. The lack 25 of an effective booster response after the third immunization of C15 was not surprising in view of previous studies where an accelerated 0, 1, and 2 month immunization schedule in baboons [Anderson et al.; J. Infect. Dis. 160:960-9 ((1989)) similarly prolonged 30 the secondary response and failed to elicit an effective tertiary booster response.

Retrospective analysis of serum and plasma from subjects C6 (Figure 1A) and C8 (Figure 1B) indicated that they became infected with HIV-1 at some point between the second and third immunizations. Serologic

evidence of HIV-1 infection was evident in the gp120 antibody assays where the titers failed to decline two weeks after the second injection and instead formed an uncharacteristic high titer plateau (Figures 1A and 1B). A similar plateau in MN-rgp120 titer after the third injection, suggested that subject C7 became infected around week 36, approximately 16 weeks after receiving the third injection (Figure 1C). Subjects C10 (Figure 1E), C11 (Figure 1D), C15 (Figure 1G), and C17 (Figure 1F) developed unexpected increases in gp120 titers, typical of HIV-1 infection, after either the third or fourth immunizations. The data obtained demonstrate that immunologic priming for MN-rgp120 antibody responses is insufficient to provide universal protection from HIV-1 infection.

5

10

15

Antibody titers to the V3 domain. To further characterize the antibody response to gp120, antibody titers were measured to a synthetic V3 domain peptide of MN-rgp120 containing the principal neutralizing 20 determinant (PND). Five of the seven subjects developed good V3 titers (1:400 to 1:4000) after the second immunization, however two subjects (C7 and C15) required three immunizations before developing significant tiers (Figures 1C and 1G). As had been 25 observed previously (11), the peak V3 titers in some individuals (e.g. C11, C10, C17) appeared to decline with each successive immunization (Figures 1D, 1E, and After HIV-1 infection, two patterns of V3 reactivity were observed. Three subjects (C6, C7, and 30 C10) showed large increases in titer to V3 domain peptides (Figures 1A, 1C, and 1E) whereas C8 (Figure 1B) showed a large decrease in V3 titer. At the time of analysis, the data were insufficient to

draw any conclusions regarding the changes in V3 titers in response to HIV-1 infection in subjects C11, C15 and C17.

The results obtained indicate that the ability to form antibodies reactive with the V3 domain at various time-points prior to HIV-1 infection is not a valid correlate of protective immunity against all strains of HIV-1.

5

25

30

35

cp4 Inhibition titers. Antibodies that block the binding of gp120 to CD4 represent a heterogeneous class of virus neutralizing antibodies. Some are known to bind to the C4 domain of gp120 [Nakamura et al.; J. Virol. 67:6179-91 (1993); Anderson et al.; J. Infect. Dis. 160:960-9 ((1989)], and some are known to recognize conformation dependent discontinuous epitopes [Berman et al.; J. Virol. 7:4464-9 (1992); Nakamura et al.; J. Virol. 67:6179-91 (1993); McKeating et al.; AIDS Research and Human Retroviruses 8:451-9 (1992); Ho et al.; J. Virol. 65:489-93 (1991);

8:451-9 (1992); Ho et al.; J. Virol. 65:489-93 (1991);
Barbas et al.; Proc. Natl. Acad. Sci. USA 91:3809-13
(1994)].
One way to detect antibodies to both types of

epitopes is to measure the ability of vaccinee sera to prevent the binding of [127]-labeled gp120 to cell surface CD4 [[Nakamura et al.; AIDS and Human Retroviruses 81875-85 (1992); Nakamura et al.; J. Virol. 67:6179-91 (1993)]. CD4 blocking titers were detected in all seven of the vaccinees prior to infection (Figure 2) with peak titers that ranged from 1:10-1:300. At the last time point prior to infection, the CD4 titers in five of the seven vaccinees was low (1:30 or less). One vaccinee (C17), however, possessed a CD4 blocking titer of about 1:300 prior to infection (Figure 2F). Thus, the lack of antibodies that block the binding of MN-rgp120 to CD4 cannot account for all

of the infections. Large increases in CD4 blocking titers (1:100-1:1,000) were seen in five of the seven subjects after HIV-1 infection. These included vaccinees C6, C7, C8, C10, and C11. These results demonstrate that the CD4 blocking titers elicited by MN-rgp120 were lower than those elicited by natural infection.

Virus neutralizing activity. The virus

10 neutralizing activity of antisera from

MN-rgp120-immunized subjects was measured using a

colorimetric assay that measured the viability of MT-4

cells after incubation with antibody treated virus

(HIV-1_{MN}). Since the actual date of infection was not

known for any of the breakthrough infections, and serum

samples were collected infrequently, the magnitude of

the neutralizing antibody response at the time of

infection is not known for any of the vaccinees.

Of the seven infections examined, the serum sample closest to the time of infection was that obtained from C7, where a neutralizing titer of 1:15 to HIV-1_{NIN} was present three weeks prior to detection of HIV-1 infection (Table 4). In all other cases, however, the interval between the last injection and the time of infection was 10 to 25 weeks.

TABLE 4

N utralization Activity of Sera from Vaccin es

Infected with HIV-1

	<u>Week</u>	<u>C6</u>	<u>C8</u>	<u>C15</u>	<u>C7</u>	<u>C11</u>	<u>C10</u>	<u>C17</u>
5	0	<10*	<10*	<10*	<10*	<10*	<10*	<10*
	2	<10	<10	<10	-	-	-	-
	4	<10*	<10*	nd*	<10*	<10*	<10*	<10*
	6	10	80	-	<10	30	150	150
	8	-	-	nd*	-	-	-	-
10	10	-	-	35	-	-	-	-
	15	_	-	-	<10	-	-	-
	16	150#	250#	-	-	30	10	<10
	24			150#	<10*	20*	<10*	<10*
	26				70	500	200	400
15	30				-	-	40	100
	33				15	-	-	-
	35				-	100	-	-
	36				30#	-	10	40
	52					30*	<10	<10
20	54					250	-	-
	57					100	-	-
	63					90	-	-
	64					-	-	<10
	77					40#	-	-
25	78						500#	10*
	80							100
	84							60

90 150

104

* - indicates immunization.

- indicates HIV-1 positive.

5

25

immunization.

nd - indicates not done.

- - indicates sample not available.

When sera from the two early infections were examined (Table 4), one individual (C6) had a peak 10 neutralizing titer of 1:10 ten weeks prior to detection of HIV-1 infection, whereas the other individual (C8) had a neutralizing titer of 1:80 ten weeks prior to detection of HIV-1 infection. Subject C15, who was immunized according to an accelerated immunization 15 schedule, developed a neutralizing titer of 1:35 after the third injection, 14 weeks prior to HIV-1 infection. Subject C10, who had a peak neutralizing titer of 1:200 following the third immunization (week 24), had no detectable titer at week 52, six months prior to the 20 first indication of HIV-1 infection (week 78).

Subject C11 possessed a neutralizing titer of 1:90 at fourteen weeks prior to detection of HIV-1 and a peak titer of 1:500 following the third immunization. Similarly vaccinee C17 had a neutralizing titer of 1:150 fourteen weeks prior to infection and a peak titer of 1:400 at two weeks after the third

Based on the rate of decay of the gp120 response of approximately two months [Belshe et al.; JAMA 272(6):475-80 (1994)], as well as the observation that neutralizing titers of 1:150 decayed to 1:10 in 10 weeks in vaccinees C10 and C17, it appears that neutralizing titers in C8, C15, C11, and C17 could have declined to 1:10 or less in the intervals between the last pre-infection serum sample and the time of HIV-1

detection.

5

10

15

20

25

The results of these studies demonstrated that all vaccinees developed some level of virus-neutralizing antibodies at some time prior to HIV-1 infection, and that the magnitude of the neutralizing response was probably low at the time of infection. In general, the magnitude of the virus-neutralizing response observed in the individuals that became infected with HIV-1 was comparable to that seen in non-infected vaccinees as described in Belshe et al.; JAMA 272(6):475-80 (1994).

sequences of Viruses. To evaluate the similarity of the breakthrough viruses with the vaccine antigen, nucleotide sequences for gp120 from all seven breakthrough viruses were determined. Envelope glycoprotein genes were amplified from proviral DNA using the polymerase chain reaction. Sequences were obtained by direct amplification of DNA from lysates of gradient-purified lymphocytes obtained directly from patient blood without any intermediate tissue culture or amplification step.

A listing of the complete gp120 sequences (two clones per specimen) is provided in Figure 3. All seven envelope glycoproteins possessed sequences typical of subtype (clade) B viruses. The overall homology with MN-rgp120 ranged from 69-80% (Table 5).

TABLE 5

Comparison of MN-rgp120 Sequence with Sequences
from Infected Vaccinees*

		MN	C6.1	C8.3	C7.2	C11.5	C10.5	C17.1	C15.2
5	MN	100	79	78	70	75	69	80	72
	C6.1		100	78	70	81	75	90	79
	C8.3			100	68	80	76	84	83
	C7.2				100	80	73	76	73
	C11.5					100	75	70	80
10	C10.5						100	70	72
	C17.1							100	87
	C15.2								100

^{* -} Data indicate percent identity.

15 Interestingly, a high percentage (four of seven) of the breakthrough viruses differed from MN-rgp120 by 25-30% [Myers et al.; Retroviruses and AIDS Database, Los Alamos National Laboratory (1992 and 1995)]. Historically this degree of sequence variation is typical of inter-subtype (intra-clade) variation rather 20 than intra-subtype variation which is expected to be in the 10-20% range [Myers et al.; Retroviruses and AIDS Database, Los Alamos National Laboratory (1992 and 19950]. Of the viruses with the greatest homology to 25 MN-rgp120, two (C6 and C8) occurred as early infections, prior to complete immunization, and one (C17) occurred as a late infection.

Polymorphism in the V3 Domain. Of particular
interest were polymorphisms in regions known to contain epitopes recognized by virus neutralizing antibodies.
The best characterized neutralizing epitope, the principal neutralizing determinant (PND), occurs at the

tip of the V3 loop. In subtype B viruses, approximately 60% possess the MN serotype-defining signature sequence, IGPGRAF (SEQ. ID. NO. 39), based on identity with the prototypic MN strain of HIV-1 [Berman et al.; J. Virol. 7:4464-9 (1992); Myers et al.; Retroviruses and AIDS Database, Los Alamos National Laboratory (1992 and 1995); La Rosa et al.; Science 249:932-5 (1990)].

5

Three of the viruses (C6, C8, and C17) possessed

the MN serotype signature sequence (Figure 3). In
contrast, four viruses possessed sequences with radical
amino acid substitutions in the PND [IGPGRAW (C7),
LGPGSTF (C11), IGPGRVL (C10), and IGPGSAF (C15)]
(SEQ. ID. NOS. 40-43, respectively), and therefore were

classified as "non-MN like" viruses. Of note, each of
the four "non-MN-like" sequences were rare (Table 6)
and were not typical of the most common "non-MN"
variants of subtype B viruses [Myers et al.;
Retroviruses and AIDS Database, Los Alamos National
Laboratory (1992 and 1995)].

TABLE 6
Frequency of Polymorphisms at the Principal
Neutralizing Determinant in HIV-1 Infected
Individuals Immunized with MN-rgp120*

5	V3 Sequence		Observed		Dataset		
	Sequence	<u>n</u>	Frequency	GNE (n=52)	LANL (n=519)	LANL.1 (n=160)	LaRosa (n=245)
	GPGRAF	3	0.42	0.67	0.57	0.66	0.60
	GPGRAW 1		0.14	0.03	0.013	0.06	0.010
10	GPGRVL	1	0.14	<0.02	0.004	<0.006	<0.008
	GPGSTF** 1	1	0.14	<0.02	<0.002	<0.006	<0.004
	GPGSAF	1	0.14	0.02	0.011	<0.006	<0.004

* - Data set GNE refers to a collection of
52 independent isolates collected in 1992;
dataset LANL refers to a collection of
519 sequences reported by Myers et al.,
Retroviruses and AIDS Database, Los Alamos

National Laboratory 1992 and 1995; LANL.1 refers
to a collection of 160 epidemiologically unlinked
individuals provided by B. Korber (personal
communication); dataset La Rosa refers to sequence
data reported by La Rosa et al., Science 249:932-5

(1990).

** - Sequences were not present in the data sets

The prevalence of viruses with PND sequences

30 matching the breakthrough viruses ranged from a high of
1.3% (C7) to a low of 0.2% (C11) in a listing of 519
subtype B sequences compiled by the Los Alamos National
Laboratory [Myers et al.; Retroviruses and AIDS
Database, Los Alamos National Laboratory (1992 and
35 1995)]. Similarly low frequencies were observed in

examined.

three other independently derived data sets (Table 6). The occurrence of these sequences did not differ significantly between data sets collected prior to 1985 [La Rosa et al.; Science 249:932-5 (1990)] and data collected 1992, or from a set of 160 epidemiologically unlinked individuals (B. Korber, personal communication). All four sets of data agreed that the prevalence of viruses with MN-like PND sequences was in the range of 60%. Based on this data, four of the seven breakthrough infections were determined to be caused by viruses that fell outside of the spectrum of viruses that the vaccine was expected to prevent.

5

10

25

30

35

Other features of breakthrough virus V3 domains.

Like MN-rgp120, the V3 domains of all of the breakthrough viruses were 36 amino acids in length. However, all seven viruses differed from MN-rgp120 with respect to the number of glycosylation sites and with respect to the syncytium-inducing (SI) signature sequence.

The sequence of MN-rgp120 is somewhat unusual [Myers et al.; Retroviruses and AIDS Database, Los Alamos National Laboratory (1992 and 1995)] in that it lacks an N-linked glycosylation site at position 306 in the V3 domain. The lack of this glycosylation site does not appear to be antigenically significant since antisera to MN-rgp120 are known to neutralize a variety of viruses (e.g. SF-2, DU6587-5, DU4489-5, CC) that possess a glycosylation site at this position [Berman et al.; J. Virol. 7:4464-9 (1992)]

In addition, the V3 domain of MN-rgp120 possessed sequence polymorphisms (R at position 311, K at position 324, K at position 328) typical of syncytium inducing viruses [Fouchier et al.; J. Virol. 66:3183-87 (1992)], whereas all seven breakthrough viruses possessed sequences associated with non-syncytium-

inducing viruses. Syncytium-inducing viruses have been associated with rapid disease progression [Tersmette et al.; J. Virol. 62:2026-32 (1988)] and T cell tropism [O'Brien et al.; Nature (London) 348:69-73 (1990); Shioda et al.; Nature (London) 349:167-9 (1991)]. To date viruses with these properties have not been recovered from any of the MN-rgp120 immunized volunteers.

5

35

Polymorphism in the V1, V2 and C4 domains.

Previous investigations have identified additional neutralizing epitopes in the V1, V2 and C4 domains of gp120 [Nakamura et al.; J. Virol. 67:6179-91 (1993); McKeating et al.; AIDS Research and Human Retroviruses 8:451-9 (1992); Ho et al.; J. Virol. 65:489-93 (1991); Barbas et al.; Proc. Natl. Acad. Sci. USA 91:3809-13 (1994); McKeating et al.; J. Virol. 67:4932-44 (1993); Moore et al.; J. Virol. 67:6136-6151 (1993); Davis et al.; J. Gen. Virol. 74:2609-17 (1993)].

The best characterized of these neutralizing epitopes is in the C4 domain which has attracted special attention because antibodies binding to this area are known to block the binding of gp120 to CD4 [Moore et al.; AIDS 3:155-63 (1989); McKeating et al.; AIDS Research and Human Retroviruses 8:451-9 (1992)]. Because the epitope is located in a conserved (C) domain, naturally-occurring polymorphism in this region is far more limited than in other neutralizing epitopes. Nakamura et al.; J. Virol. 67:6179-91 (1993) reported that the binding of a number of neutralizing MAbs was dependent on K at position 429.

Comparison of the sequence of MN-rgp120 with other strains of HIV-1 showed that a common polymorphism, involving the substitution of E for K, occurs at this position. Indeed, substrains of the same virus isolate often show polymorphism at this position. The HXB2

substrain of HIV-1_{IAI} contains K at position 429, whereas the BH10, IIIB, and LAV substrains of the HIV-1_{IAI} contain E at this position [Nakamura et al.; J. Virol. 67:6179-91 (1993)]. Similarly, the 1984 isolate of HIV-1_{MN} exhibited E at this position, while the 1990 isolate of HIV-1_{MN}, used to produce MN-rgp120, possessed K at this position.

5

10

15

20

25

30

35

When the sequences of the infected vaccine recipients were examined (Figure 3), the virus from subject C17, like MN-rgp120 contained K at position 429, whereas the six other viruses that differed from the vaccine immunogen possessed E at this position. These results demonstrated that six of the seven breakthrough viruses differed from the vaccine immunogen at the CD4-blocking, neutralizing epitope in the C4 domain of gp120.

Studies with monoclonal antibodies have defined neutralizing epitopes in the V1 and V2 domains of gp120 [McKeating et al.; J. Virol. 67:4932-44 (1993); Moore et al.; J. Virol. 67:6136-6151 (1993); Davis et al.; J. Gen. Virol. 74:2609-17 (1993)]. Like the polymorphisms that occur in the C4 domain, the V2 domains exhibit several common polymorphisms that affect the binding of virus neutralizing antibodies. One such polymorphism occurs at position 171 which is critically important for the binding of murine MAb 1025, whereas residue 187 is important for the binding of MAb several MAbs represented by 1088.

When the V2 domain sequences were examined (Figure 3), all of the infected-vaccinee viruses differed from MN-rgp120 in that R replaced G at position 171 and I or V replaced E at position 187. Antibodies recognizing these adjacent sites in the V2 domain of MN-rgp120 would not be expected to neutralize viruses with radical amino acid substitutions at these position. Thus, all seven

breakthrough viruses differed from MN-rgp120 at a neutralizing epitope in the V2 domain of gp120.

Other neutralizing epitopes have been reported in the V1 domain of gp120 [O'Brien et al.; Nature (London)

348:69-73 (1990); McKeating et al.; J. Virol.

67:4932-44 (1993)]. Although the neutralizing epitopes in the V1 domain of MN-rgp120 have not been characterized, the polymorphism seen among the breakthrough viruses in this region was interesting.

Particularly striking (Figure 3) was that the length of this domain ranged from 20 amino acids (C17) to 45 amino acids (C6), and the number of N-linked glycosylation sites ranged from 2 to 6

45 amino acids (C6), and the number of N-linked glycosylation sites ranged from 2 to 6. In contrast, the V1 domain of MN-rgp120 is 31 amino acids in length and encodes three N-linked glycosylation sites.

Although examination of sequence databases suggest that variation in the V2 region is comparable to the V1 region, the V2 region of the breakthrough viruses showed less variation than expected. Specifically, the length of the V2 region ranged from 36 amino acids (C7) to 39 amino acids in length, with six of seven viruses containing three N-linked glycosylation sites in this domain. A high degree of polymorphism was found in the V4 region where sequences ranged from 26 (C10) to 33 (C15, C7) amino acids in length and contained either 4 or 5 N-linked glycosylation sites.

20

25

30

35

Antigenicity of envelope glycoproteins from breakthrough viruses. To determine the significance of sequence variation on glycoprotein antigenicity, recombinant gp120 was prepared from the viruses of all seven infected vaccinees (Figure 4). In these studies a series of MAbs was assembled and their binding to MN-rgp120 was compared to that of rgp120 from the vaccinee isolates by ELISA (Table 7).

TABLE 7
Relative Reactivity of MAb Binding to rgp120 fr m
Infected Subj cts Compared with Binding to MN-rgp120

			<u>v3</u>	Discont	inuous	<u>C8</u> <u>V</u>		
5	gp120	1034	<u>50.1</u>	<u>1.5E</u>	1025	1024	1088	
	MN	1.0	1.00	1.00	1.00	1.00	1.00	
	C6.1 C6.5	0.37 0.33	0.37 0.33	0.17 0.75	0.00	0.00	0.00	
10	C8.3 C8.6	0.11 0.14	0.37 0.34	0.38 0.29	0.00	0.00	0.00	
	C7.2	0.47	0.60	0.71	0.00	0.00	0.00	
	C11.5 C11.7	0.00	0.00 0.00	0.17 0.17	0.00	0.00	0.00	
15	C10.5 C10.7	0.33 0.42	0.40 0.48	0.46 0.50	0.24	0.03 0.07	0.04 0.09	
	C17.1 C17.3	0.33 0.37	0.52 0.56	0.33	0.00	0.30	0.07 0.06	
	C15.2 C15.3	0.00	0.47 0.37	0.92 0.63	0.00 0.00	0.00	0.00	

20

25

30

35

40

* - Relative reactivity values represent ratio of optical densities obtained with rgp120 from patient isolates divided by optical density obtained for MN-rgp120 at a MAb concentration of 2 micrograms per milliliter.

In control experiments, the binding of MAb 5B6 (which is specific for the HSV gD-1 flag epitope fused to the N terminus of all of the rgpl20 protein) was used to standardize the amount of gpl20 from each isolate (Figure 5A). These studies demonstrated that the assay was carried out under conditions where equivalent amount of rgpl20s were captured onto wells of microtiter plates.

The antigenic structure of the V3 domain was examined using the 1034 MAb (isolated from mice immunized with MN-rgp120 as described in Nakamura et al.; J. Virol. 67:6179-91 (1993) and the 50.1 MAb (prepared from mice immunized with a synthetic V3 domain peptide as described in Rini et al.; Proc.

Natl. Acad. Sci. USA 90:6325-9 (1993). Both MAbs are known to exhibit potent virus neutralizing activity. When binding to the recombinant proteins was examined, the MAb binding to MN-rgp120 was at least 10-fold greater than to any of the breakthrough virus envelope proteins (Figure 5 B and C). Surprisingly, rgp120 from the three patient isolates (C8, C6, and C17) that possessed the MN serotype-defining sequence, IGPGRAF (SEQ. ID. No. 39), varied from one another in their MAb binding activity. Thus, the binding of MAb 1034 and MAb 50.1 to rgp120 from C17 was significantly greater than the binding to rgp120s from C6 and C8.

5

10

A distinction in the epitopes recognized by these MAbs was evident since C6-rgp120 and C8-rgp120 gave comparable binding with 50.1, whereas 1034 bound better 15 to the C6-derived protein than the C8-derived protein. The poorest MAb reactivity was with rgpl20s from C11 and C15. This result was consistent with sequence analysis demonstrating that these two viruses both possessed the radical substitution of S for R at 20 position 18 in the V3 domain. Surprisingly, both of these MAbs exhibited better than expected binding to rgp120 from the C7 and C10 viruses. Like MN-rgp120, both proteins contained the penta-peptide, IGPGR sequence (SEQ. ID. NO. 44) in the V3 loop, but differed 25 from MN-rgp120 in that V and L replaced A and F at positions 319 and 320 in gp120 from C10, and W replaced F at position 320 in gp120 from C7. These results indicate that R at position 318 is essential for the binding of these two MAbs, and that the epitopes 30 recognized by 1034 and 50.1 are not completely destroyed by the hydrophobic substitutions at positions 319 and 320.

As predicted from the sequence data, there was

little if any binding to the breakthrough virus rgp120s
using MAbs (1088 and 1025) directed to the V2 region of

MN-rgp120. Also consistent with sequence data was the observation that MAb 1024 directed to the C4 domain of MN-rgp120 gave some reactivity with C17-rgp120 which, like MN-rgp120 contained K at position 429, but gave no reactivity with the other isolates that contained E at residue 429.

Together, these studies demonstrated that the antigenic structure of all seven breakthrough viruses differed from the vaccine immunogen at three well characterized neutralizing epitopes.

5

10

15

A totally different pattern of reactivity was observed with the human hybridoma, MAb 15e, prepared from an HIV-1 infected individual as described in Ho et al.; J. Virol. 65:489-93 (1991). With this MAb, the greatest binding was achieved with MN-rgp120 and rgp120 from C7, and the poorest reactivity was seen with the two clones of rgp120 from the C11. Moderate, but comparable reactivity was seen with rgp120s from the C10 and C17.

These results demonstrate that the 15e epitope is 20 polymorphic, and that the epitope is conserved on MN-rgp120 and rgp120 from C7, but has been lost on Interestingly, the two different rgp120s from C11. clones of gp120 derived from C6 gave strikingly different patterns of antibody binding. Thus, rgp120 25 from clone C6.5 exhibited strong reactivity with this antibody, whereas rgp120 from clones C6.1 exhibited significantly weaker activity with this MAb. Comparison of sequence data (Figure 3) showed that the two C6 clones differed at 6 amino acid positions. 30 Based on comparative binding to the other viral proteins of known sequence, it appeared that the substitution of K for I at position 351 in the C3 domain of gp120 could account for the difference in binding activity. This result is also consistent with 35 both clones of C11 similarly containing a positively-

charged K at this position, and also being poorly reactive with this MAb. Alternatively, a T for I substitution at position 439 in the C4 domain could account for the difference in 15e binding between C6.1 and C6.5. Although the inability of the two C11 clones to bind 15e cannot be explained by polymorphism at this position in the C4 domain, they could be affected by the adjacent T for M substitution at position 434.

10 Discussion

In these studies, the viruses and immune responses in seven of nine vaccinees who became infected with HIV-1 through high risk activity while participating in Phase I or Phase 2 trials of MN-rgp120, a candidate

HIV-1 vaccine were analyzed. Such infections would be expected to occur for one of two reasons: 1) lack of sufficient immune response at the time of infection; or 2) infection with viruses that fall outside of the antigenic spectrum expected to be covered by the vaccine immunogen. The data indicate that both explanations may be involved with the infections observed (Table 8).

TABLE 8
Summary of Breakthrough Infecti ns

Homologous to MN-rqp120

			MN-rgp120			
5	Case No.	Adequate	Homology	V3	C4	V2
		Immunization	(8)	PND	Epitope	Epitope
	C6		79	+	-	-
	C8	-	78	+	-	-
	C15	-	72	-	-	-
	C7	-	70	-	-	-
10	C11	+	75	-	-	-
	C10	+	69	-	-	-
	C17	+	80	+	+	-

15 Two of the infections occurred in individuals who failed to receive the minimum three doses of vaccine typically required for the induction of protective immunity with protein subunit vaccines (e.g. hepatitis B virus formulated in alum adjuvant as described in Francis et al.; Ann. Int. Med. 97:362-6 (1982). 20 additional breakthrough infections occurred in vaccinees who had weak or undetectable primary (C7) and booster (C15) responses. Of the three individuals who became infected with HIV-1 after receiving three or 25 more productive immunizations (C10, C11, and C17), at least two, and possibly all three, appear to have become infected more than six months after receiving their last immunization. Because antibody titers to MN-rgp120 typically decay with a half-time of 2 to 2.5 months [Belshe et al.; JAMA 272(6):475-80 (1994); 30 Berman et al.; AIDS 8:591-601 (1994)], antibody titers would be expected to have decayed at least eight-fold and possibly as much as sixty four-fold at the time of infection. Thus, the lack of a sufficient immune

response at the time of infection represents a potential explanation for at least six of the seven breakthrough infections.

Data from vaccine efficacy studies in gp160 immunized chimpanzees [McElrath et al.; Longitudinal Vaccine-Induced Immunity and Risk Behavior of Study Participants in AVEG Phase II Protocol 201. In: Abstracts from Eighth Annual Meeting of the National Cooperative Vaccine Development Groups for AIDS. Bethseda, MD 1996:216] challenged with HIV-1, and gp120-immunized rhesus macaques challenged with a chimeric SIV/HIV-1 virus (SHIV) suggest that the magnitude of the neutralizing antibody response at the time of infection is a critical correlate of protective immunity. If maintaining neutralizing antibody titers proves to be a valid correlate of protective immunity in humans, then formulations (e.g. novel adjuvants) or immunization regimes (frequent boosting) designed to maximize the antibody responses may be required to achieve long lasting protection. Use of a booster

10

15

20

25

30

35

The other likely explanation for the late infections is the antigenic difference between the vaccine and the breakthrough virus envelope glycoproteins. This explanation is supported by the observation that four of the seven breakthrough viruses possessed envelope glycoproteins that differed from the MN-rgp120 by 25-30% at the amino acid level. Differences of this magnitude have historically

every six months may be advantageous.

[Myers et al.; Retroviruses and AIDS Database, Los Alamos National Laboratory (1992 and 1995)] been associated with inter-subtype variation and far exceeds the average 10-20% variation expected for viruses within the same subtype.

Although the biologic significance of sequence variation in many regions of the envelope glycoprotein

is unclear, polymorphism at neutralizing epitopes is an important factor that affects vaccine efficacy. Previous studies [Salmon-Ceron et al.; AIDS Res. and Human Retroviruses 11:1479-86 (1995); Javaherian et al.; Science 250:1590-3 (1990)] have demonstrated that the breadth of neutralizing activity that could be elicited by HIV-1 envelope derived vaccines was critically dependent on the sequence of epitopes in the V3 domain (e.g.; the PND). Thus, candidate vaccines based on the LAI strain of HIV-1 (the prototypic "non-MN-like" subtype B virus), exhibited little or no cross neutralizing activity with subtype B viruses, whereas vaccines that contained the "MN-like-" PND sequence (IGPGRAF) (SEQ. ID. NO. 44) exhibited broad cross neutralizing activity. That four of the seven breakthrough viruses possessed envelope glycoproteins with radical amino acid substitutions in the PND is consistent with the explanation that differences in antigenic structure explain some of these infections.

10

15

20

25

30

35

Over the last few years, it has become clear that polymorphism among "MN-like" viruses occurs at neutralizing epitopes outside of the PND. The best example occurs in the C4 domain where two antigenically distinct variants are distinguished by the presence of either K or E at position 429 [Moore et al.; AIDS 3:155-63 (1989)]. Because six of the seven breakthrough viruses differed from the vaccine strain in that they contained E rather than K at position 429, antibodies raised to the C4 domain of MN-rgp120 were unlikely to neutralize the viruses infecting in six of the seven vaccinees.

Other neutralizing epitopes are known to be present in the V1 and V2 domains of gp120. Although these regions are highly variable, due to insertions and deletions, neutralizing epitopes have been described by McKeating et al.; J. Virol. 67:4932-44

(1993); Moore et al.; J. Virol. 67:6136-6151 (1993); and Davis et al.; J. Gen. Virol. 74:2609-17 (1993). Several of these epitopes overlap an amino terminal sequence of the V2 domain containing the tri-peptide sequence RDK at positions corresponding to 142 to 144 of MN-rgp120 [McKeating et al.; J. Virol. 67:4932-44 (1993); Moore et al.; J. Virol. 67:6136-6151 (1993)]. Like the C4 epitope, variation in this sequence is known to occur between different substrains derived from the same parental isolate. Since all seven breakthrough viruses differed from MN-rgp120 in that they possessed the RDK sequence, rather than the GDK sequence present in the vaccine antigen, neutralizing antibodies to the V2 domain of MN-rgp120 would not have been expected neutralize any of the viruses recovered from the vaccinees immunized with MN-rgp120.

5

10

15

20

25

30

35

Although polymorphisms at neutralizing epitopes might account for the lack of protection in most of the infections, this does not appear to explain the infection of vaccinee C17, who was infected by a virus that matched MN-rgp120 in the V3 and C4 domains. If a difference in sequence was responsible for the lack of protection in this case, the critical difference might relate to the unusual sequence in the V1 domain of gp120 from this breakthrough virus. Several studies have shown that the V1 domain possesses epitopes recognized by virus neutralizing monoclonal antibodies [McKeating et al.; J. Virol. 67:4932-44 (1993); Davis et al.; J. Gen. Virol. 74:2609-17 (1993); Kayman et al.; J. Virol. 68:400-410 (1994)].

Although far less is known about the V1 epitopes relative to other neutralizing sites, the V1 epitopes appear to be conformation-dependent, and antisera from HIV-1 infected individuals recognize epitopes in the V1 and V2 domains [McKeating et al.; J. Virol. 67:4932-44 (1993); Kayman et al.; J. Virol. 68:400-410 (1994)].

The V1 sequence of the virus from C17 is noteworthy because it is smaller and contains fewer N-linked glycosylation sites than that of MN-rgp120 or any of the other breakthrough viruses. By the same token, the envelope glycoproteins from C11 and C6 are noteworthy because they are significantly larger and contain more glycosylation sites than MN-rgp120 or the other breakthrough viruses.

5

10

15

20

25

30

35

While differences in amino acid sequence can provide clues to differences in antigenic structure, the consequences of such polymorphism can only be proven through antibody binding studies. To correlate differences in sequence with differences in antigenic structure, gp120 from two clones each of all seven breakthrough viruses was expressed and the antigenicity of the clones with a panel of monoclonal antibodies was examined. As predicted from the sequence data, none of the breakthrough virus envelope glycoproteins reacted with neutralizing MAbs to the V2 domain of MN-rgp120. When MAbs to the C4 domain were examined, only the C17 envelope glycoprotein (that matched MN-rgp120 with respect to K429) showed significant, albeit lower, Surprisingly, the three breakthrough envelope binding. glycoproteins that contained the subtype B PND consensus sequence, IGPGRAF (SEQ. ID. NO. !!), gave poor reactivity with all three PND directed MAbs, even though they possessed PND sequences closely related to the vaccine immunogen. Thus, all three of the vaccinee isolates appeared to possess changes outside of the recognition site that interfered with MAb binding.

It has been known for many years that resistance to neutralization in vitro can sometimes be attributed to mutations in remote sequences that alter the conformation of neutralizing epitopes and interfere with recognition by virus neutralizing antibodies [Nara et al.; J. Virol. 64:3779-91 (1990); Cordonnier

et al.; Nature 340:571-4 (1989)]. Together, these results indicate that the antigenic structure of the envelope glycoproteins recovered from the breakthrough viruses differed significantly from that of the vaccine antigen.

5

10

30

35

A novel result was the localization of residues in the C3 domain that appeared to affect the binding of the virus neutralizing human MAb, 15e. This MAb is known to recognize a discontinuous epitope, block CD4 binding, and neutralize a variety of laboratory and primary isolates of HIV-1 [Ho et al.; J. Virol. 65:489-93 (1991); Thali et al.; J. Virol. 66:5635-5641 (1992); Moore et al.; AIDS Res. Hum. Retroviruses 9:1179-1187 (1993)].

Comparative binding to envelope glycoproteins from 15 the breakthrough viruses indicated that recognition by this antibody is critically dependent on residues in the C3 or C4 domains of gp120. The unique occurrence of a positively charged K at position 351 in the C3 domain provides a common explanation for the inability 20 of the C11.5, C11.7 and C6.1 strains of HIV-1 to bind to 15e. Alternatively, it is possible that different amino acid substitutions in different locations account for the failure of 15e to bind to rgp120s from the C6 and C11 clones. The only obvious positions where 25 substitutions of this type occur are in the C4 domain where T replaces M at 434 (C11) and T replaces I at 439.

The present studies demonstrate that the current formulation of MN-rgp120 is less than 100% effective against HIV-1 infection. Based on previous in vitro and in vivo studies with MN-rgp120, protection from natural HIV-1 infection in humans is expected to depend on a threshold concentration of virus-neutralizing antibodies, and antigenic similarity between the vaccine immunogen and the challenge virus.

In this regard, only one of the seven breakthrough infections (C17) was unexpected. This individual received a full course of immunizations yet became infected with a virus similar to MN-rgp120 at at least two important neutralizing epitopes (V3 and C4 domains). This infection might be related to the magnitude of the antibody response at the time of infection, or antigenic differences between the breakthrough virus and the vaccine strain, or circumstances of infection (e.g., ulcerative lesions, infection by donor with acute infection or high viremia), not monitored in this protocol. Alternatively this individual may represent a true vaccine failure, without clear explanation.

5

10

15

20

25

30

35

On balance, the analysis of breakthrough infections described herein did not uncover any data that would discourage the continued development of MN-rgp120 as a vaccine to prevent HIV-1 infection. The results support speculation that enhancing vaccine immunogenicity (as by additional booster immunizations) may be required to maintain long term protective immunity, and that the addition of rgp120 from other antigenically different strains of virus in addition to MN-rgp120 are useful to expand the breadth of protection.

The availability of viruses and viral glycoproteins derived from breakthrough infections may provide an important means to streamline the process of identifying new antigens for inclusion into a multivalent vaccine. Recombinant viral glycoproteins prepared from breakthrough viruses, by definition, possess antigenic structures that are significantly different from MN-rgp120, and are be representative of viruses currently being transmitted. Thus, combining rgp120 from breakthrough viruses with MN-rgp120 is an effective way complement and significantly expand

antigenic complexity and increase breadth of cross neutralizing activity.

SEQUENCE LISTING

```
(1) GENERAL INFORMATION:
           (i) APPLICANT: Berman, Phillip W.
          (ii) TITLE OF INVENTION: HIV ENVELOPE POLYPEPTIDES AND
 5
                  VACCINE
         (iii) NUMBER OF SEQUENCES: 44
          (iv) CORRESPONDENCE ADDRESS:
                 (A) ADDRESSEE: SKJERVEN, MORRILL, MACPHERSON, ET AL.
                 (B) STREET: 25 Metro Drive, Suite 700
10
                 (C) CITY: San Jose
                 (D) STATE: California
                 (E) COUNTRY: USA
                 (F) ZIP: 95110
15
           (v) COMPUTER READABLE FORM:
                (A) MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
                (B) COMPUTER: IBM PC compatible
                (C) OPERATING SYSTEM: PC-DOS/MS-DOS
                (D) SOFTWARE: WinPatin (Genentech)
          (vi) CURRENT APPLICATION DATA:
20
                (A) APPLICATION NUMBER:
                 (B) FILING DATE:
                (C) CLASSIFICATION:
        (viii) ATTORNEY/AGENT INFORMATION:
                 (A) NAME: Terlizzi, Laura
25
                 (B) REGISTRATION NUMBER: 31,307
                 (C) REFERENCE/DOCKET NUMBER: M-3897 US
          (ix) TELECOMMUNICATION INFORMATION:
                (A) TELEPHONE: (408) 453-9200
(B) TELEFAX: (408) 453-7979
30
      (2) INFORMATION FOR SEQ ID NO:1:
         (i) SEQUENCE CHARACTERISTICS:
             (A) LENGTH: 1503 base pairs
             (B) TYPE: Nucleic Acid
35
             (C) STRANDEDNESS: Single
             (D) TOPOLOGY: Linear
        (xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:
           GGG GTA CCT GTG TGG AAG GAA GCA ACC ACC ACT CTA 36
40
           Gly Val Pro Val Trp Lys Glu Ala Thr Thr Leu
       TTT TGT GCA TCA GAT GCT AAA GCA TAT GAC ACA GAG GTG 75
       Phe Cys Ala Ser Asp Ala Lys Ala Tyr Asp Thr Glu Val
45
       CAT AAT GTT TGG GCC ACA CAT GCT TGT GTA CCC ACA GAC 114
       His Asn Val Trp Ala Thr His Ala Cys Val Pro Thr Asp
50
                        30
       CCA AAC CCA CAA GAA ATG GTA TTG GAA AAT GTG ACA GAA 153
       Pro Asn Pro Gln Glu Met Val Leu Glu Asn Val Thr Glu
55
       GAT TTT AAC ATG TGG AAA AAT GAC ATG GTA GAA CAG ATG 192
       Asp Phe Asn Met Trp Lys Asn Asp Met Val Glu Gln Met
60
      CAT GAG GAT ATA ATC AGT TTA TGG GAT CAA AGC CTA AAA 231
       His Glu Asp Ile Ile Ser Leu Trp Asp Gln Ser Leu Lys
                            70
```

	CCA Pro	TG:	GTA Val 80	. Lys	TTA Leu	ACC Thr	CCF Pro	CTC Let	2 Cy	T AT:	r AC:	î TT r Le	'AA A u As: 9	
5	TGC Cys	ACC Thi	AAT Asn	TGG	AAG Lys 95	Lys	TAA i	GAT Asp	r AC:	100	Th	AA 1 As	T AG: n Se:	r 309
10	AGT Ser	AGT Ser 105	Thr	ACA Thr	ACT Thr	AAT Asn	AAT Asn 110	Ser	AG1	C GCT	ACA Thi	A GC' Ala	a Asr	348
15	AGT Ser	AG1 Ser	AGT Ser	ACT Thr 120	Thr	ACT Thr	AAT Asn	AGT Ser	Ser 125	Trp	GG# Gly	GAC Glu	G ATA	387
20	AAG Lys 130	Glu	GGA Gly	GAA Glu	ATA Ile	AAG Lys 135	AAC Asn	TGC	TCT	TTC Phe	AAT Asn 140	Ile	ACC Thr	426
	ACA Thr	AGC Ser	ATA Ile 145	AGA Arg	GAC Asp	AAG Lys	GTG Val	AAG Lys 150	Lys	GAA Glu	TAT Tyr	GCA Ala	CTT Leu 155	
25	TTT Phe	TAT Tyr	AGC Ser	CTT Leu	GAT Asp 160	GTA Val	GTA Val	CCA Pro	ATA Ile	GAA Glu 165	AAT Asn	CAT Asp	AAT Asn	504
30	ACT Thr	AGC Ser 170	TAT Tyr	AGG Arg	TTG Leu	AGA Arg	AGT Ser 175	TGT Cys	AAC Asn	ACC Thr	TCA Ser	CTC Val 180	Ile	543
35	ACA Thr	CAA Gln	GCC Ala	TGT Cys 185	CCA Pro	AAG Lys	GTA Val	ACT Thr	TTT Phe 190	GAG Glu	CCA Pro	ATT	CCC Pro	582
40	ATA Ile 195	CAT His	TAT Tyr	TGT Cys	ACC Thr	CCG Pro 200	GCT Ala	GGT Gly	TTT Phe	GCG Ala	ATT Ile 205	CTG Leu	AAG Lys	621
••	TGT Cys	AGA Arg	GAT Asp 210	AAA Lys	AAG Lys	TTC Phe	AAT Asn	GGA Gly 215	ACA Thr	GGA Gly	CCA Pro	TCC Cys	AAA Lys 220	660
45	AAT Asn	GTT Val	AGC Ser	ACA Thr	GTA Val 225	CAA Gln	TGT Cys	GCA Ala	CAT His	GGA Gly 230	ATT Ile	AAG Lys	CCA Pro	699
50	GTA Val	GTG Val 235	TCA Ser	Thr	CAA Gln	Leu	Leu	Leu	AAT Asn	GGC Gly	AGC Ser	CTA Leu 245	GCA Ala	738
55	GAA Glu	GAA Glu	GAG Glu	GTA Val 250	ATA Ile	ATT Ile	AGA Arg	TCT Ser	GCC Ala 255	AAT Asn	TTC Phe	TCA Ser	AAC Asn	777
60	AAT Asn 260	GCT Ala	AAA Lys	ATC Ile	ATA Ile	ATA Ile 265	GTA Val	CAG Gln	TTG Leu	AGG Arg	GAA Glu 270	CCT Pro	GTA Val	816
	GAA Glu	ATT Ile	AAT Asn 275	TGT Cys	ACA Thr	AGA Arg	Pro	AGC Ser 280	AAC Asn	AAT Asn	ACA Thr	ATA 11e	AAA Lys 285	855

	GGT Gly	ATA Ile	CAC His	ATA Ile	GGA Gly 290	CCA Pro	GGG Gly	AGA Arg	GCA Ala	TTT Phe 295	TAT Tyr	GCA Ala	ACA Thr	894
5	GGA Gly	GAC Asp 300	ATA Ile	CGA Arg	GGA Gly	GAT Asp	ATA Ile 305	AGA Arg	CAA Gln	GCA Ala	CAT His	TGT Cys 310	AAC Asn	933
10	ATT Ile	AGT Ser	GGA Gly	GCA Ala 315	AAA Lys	TGG Trp	TAA Asn	AAC Asn	ACT Thr 320	TTA Leu	AAG Lys	AAG Lys	GTA Val	972
15	GTT Val 325	AAA Lys	AAA Lys	TTA Leu	Lys	GAA Glu 330	CAA Gln	TTT Phe	CCA Pro	AAT Asn	AAA Lys 335	ACA Thr	ATA Ile	1011
••	GTC Val	TTT Phe	AAC Asn 340	CAT His	TCC Ser	TCA Ser	GGA Gly	GGG Gly 345	GAC Asp	CCA Pro	GAA Glu	ATT Ile	GTA Val 350	1050
20	ATG Met	CAC His	AGT Ser	TTT Phe	AAT Asn 355	TGT Cys	CAA Gln	GGG Gly	GAA Glu	TTT Phe 360	TTC Phe	TAC Tyr	TGT Cys	1089
25	AAT Asn	ACA Thr 365	ACA Thr	AAG Lys	CTG Leu	TTT Phe	AAT Asn 370	AGT Ser	ACT Thr	TGG Trp	AAT Asn	GAT Asp 375	ACT Thr	1128
30	ACA Thr	GAG Glu	TCA Ser	AAT Asn 380	AAC Asn	AAT Asn	GAT Asp	AGT Ser	ACT Thr 385	ATT Ile	ACA Thr	CTC Leu	CCA Pro	1167
35	TGC Cys 390	AGA Arg	ATA Ile	AAA Lys	CAA Gln	ATT Ile 395	ATA Ile	AAC Asn	ATG Met	TGG Trp	CAG Gln 400	GAA Glu	ATA Ile	1206
	GGA Gly	AAA Lys	GCA Ala 405	ATG Met	TAT Tyr	GCC Ala	CCT Pro	CCC Pro 410	ACC Thr	AGA Arg	GGA Gly	GAA Glu	ATT Ile 415	1245
40	AAA Lys	TGT Cys	TCA Ser	TCA Ser	AAT Asn 420	ATT Ile	ACA Thr	GGA Gly	CTA Leu	CTG Leu 425	TTA Leu	ATA Ile	AGA Arg	1284
45	GAT Asp	GGT Gly 430	GGT Gly	ATT Ile	AAC Asn	ACT Thr	AGC Ser 435	GAT Asp	GCC Ala	ACC Thr	GAG Glu	ACC Thr 440	TTC Phe	1323
50	AGA Arg	CCG Pro	GGA Gly	GGA Gly 445	GGA Gly	GAT Asp	ATG Met	AGG Arg	GAC Asp 450	AAT Asn	TGG Trp	AGA Arg	AGT Ser	1362
55	GAA Glu 455	TTA Leu	TAT Tyr	AAA Lys	TAT Tyr	AAA Lys 064	GTA Val	GTG Val	AAA Lys	ATT Ile	GAG Glu 465	CCA Pro	TTA Leu	1401
	GGA Gly	GTA Val	GCA Ala 470	CCC Pro	ACC Thr	AAG Lys	GCA Ala	AAG Lys 475	AGA Arg	AGA Arg	GTG Val	GTG Val	CAG Gln 480	1440
60	AGA Arg	GAA Glu	AAA Lys	AGA Arg	GCA Ala 485	GTA Val	ACA Thr	CTA Leu	GGA Gly	GCT Ala 490	ATG Met	TTC Phe	CTT Leu	1479

GGG TTC TTA GGA GCA TAA AGC TTC 1503 Gly Phe Leu Gly Ala Xaa Ser Phe 495 500 501

5	5 (2) INFORMATION FOR SEQ ID NO:2: (i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 501 amino acids (B) TYPE: Amino Acid (D) TOPOLOGY: Linear													
10		DESCRIPTION:	SEQ ID NO:2	:										
	Gly Val Pro Va 1	l Trp Lys Glo	Ala Thr Thr	Thr Leu Phe Cys Ala										
15	Ser Asp Ala Ly	s Ala Tyr Asp 20	Thr Glu Val 25	His Asn Val Trp Ala 30										
20		35	40	Pro Gln Glu Met Val 45										
		50	55	Trp Lys Asn Asp Met 60										
25		65	70	Leu Trp Asp Gln Ser 75										
2.0		80	85	Cys Ile Thr Leu Asn 90										
30		95	100	Thr Asn Ser Ser Ser 105										
35		110	115	Asn Ser Ser Ser Thr 120										
		125	130	Glu Gly Glu Ile Lys 135										
40		140	145	Arg Asp Lys Val Lys 150										
4 E		155	160	/al Val Pro Ile Glu 165										
45		170	175	Cys Asn Thr Ser Val 180										
50		185	190	lu Pro Ile Pro Ile 195										
		200	205	eu Lys Cys Arg Asp 210										
55		215	220	sn Val Ser Thr Val 225										
60		230	235	er Thr Gln Leu Leu 240										
60		245	250	le Ile Arg Ser Ala 255										
65	Asn Phe Ser Asn	Asn Ala Lys 260	Ile Ile Ile V 265	al Gln Leu Arg Glu 270										

```
Pro Val Glu Ile Asn Cys Thr Arg Pro Ser Asn Asn Thr Ile Lys
                                                                 285
                       275
                                            280
       Gly Ile His Ile Gly Pro Gly Arg Ala Phe Tyr Ala Thr Gly Asp
                                            295
                                                                300
 5
                       290
       Ile Arg Gly Asp Ile Arg Gln Ala His Cys Asn Ile Ser Gly Ala
                                            310
       Lys Trp Asn Asn Thr Leu Lys Lys Val Val Lys Lys Leu Lys Glu
10
                                            325
                       320
       Gln Phe Pro Asn Lys Thr Ile Val Phe Asn His Ser Ser Gly Gly
                                            340
                       335
15
       Asp Pro Glu Ile Val Met His Ser Phe Asn Cys Gln Gly Glu Phe
                                            355
                       350
       Phe Tyr Cys Asn Thr Thr Lys Leu Phe Asn Ser Thr Trp Asn Asp
                                            370
20
                       365
       Thr Thr Glu Ser Asn Asn Asn Ser Thr Ile Thr Leu Pro Cys
                       380
                                            385
       Arg Ile Lys Gln Ile Ile Asn Met Trp Gln Glu Ile Gly Lys Ala
25
                                            400
                                                                405
       Met Tyr Ala Pro Pro Thr Arg Gly Glu Ile Lys Cys Ser Ser Asn
                       410
30
       Ile Thr Gly Leu Leu Ile Arg Asp Gly Gly Ile Asn Thr Ser
                       425
       Asp Ala Thr Glu Thr Phe Arg Pro Gly Gly Gly Asp Met Arg Asp
35
                       440
       Asn Trp Arg Ser Glu Leu Tyr Lys Tyr Lys Val Val Lys Ile Glu
                                           460
                       455
       Pro Leu Gly Val Ala Pro Thr Lys Ala Lys Arg Arg Val Val Gln
40
                                           475
       Arg Glu Lys Arg Ala Val Thr Leu Gly Ala Met Phe Leu Gly Phe
                       485
                                           490
45
       Leu Gly Ala Xaa Ser Phe
                       500 501
      (2) INFORMATION FOR SEQ ID NO:3:
         (i) SEQUENCE CHARACTERISTICS:
50
             (A) LENGTH: 1503 base pairs
             (B) TYPE: Nucleic Acid
             (C) STRANDEDNESS: Single
             (D) TOPOLOGY: Linear
        (xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:
55
           GGG GTA CCT GTA TGG AAA GAA GCA ACC ACC ACT CTA 36
           Gly Val Pro Val Trp Lys Glu Ala Thr Thr Thr Leu
60
       TTT TGT GCA TCA GAT GCT AAA GCA TAT GAC ACA GAG GTG 75
       Phe Cys Ala Ser Asp Ala Lys Ala Tyr Asp Thr Glu Val
                15
```

	CA Hi	T AA S As	AT GI in Va	T TG	G GCG p Ala 30	a Th	A CAS	r GC 3 Al	T TG: a Cy:	r GT	Pro	ACA Thr	. Yal	2 114
5	CC. Pr	O MS	c cc n Pr 0	A CAR	A GA# n Glu	A ATO	G GTA Val	Le	G GA/ u Glu	A AAT	r GTG n Val	ACA Thr 50	Glu	153
10	GA' As	T TT p Ph	T AA e As	C ATO n Met 55	rrp	AAA Lys	AAT Asn	, GV	C ATO P Met 60	. val	GAA Glu	CAG Gln	ATG Met	192
15	65	9 GT	u xa	T ATA a Ile	. Ile	70	Leu	Tr	o Asp	Gln	Ser 75	Leu	Lys	
20	PFC	o Cy:	s va. 8		Leu	Thr	Pro	Leu 85	Cys	Ile	Thr	Leu	Asn 90	
	Cys	Thi	C ASI	TGG Trp	Lys 95	Glu	Asn	Asp	Thr	Lys 100	Thr	Asn	Ser	
25	ser	105	Thi	ACA Thr	Thr	Asn	Asn 110	Ser	Ser	Ala	Thr	Ala 115	Asn	
30	Ser	Ser	Ser	Thr 120	Thr	Thr	Asn	Ser	Ser 125	Trp	Gly	Glu	Ile	
35	130	GIU	GIA	GAA Glu	He	Lys 135	Asn	Cys	Ser	Phe	Asn 140	Ile	Thr	
40	The	GIÀ	11e 145	AGA Arg	Asp	Lys	Val	Lys 150	Lys	Glu	Tyr	Ala.	Leu 155	
	Pne	Tyr	Ser	CTT Leu	160	Val	Val	Pro	Ile	Glu 165	Asn /	Asp A	Asn	
45	Thr	170	Tyr	AGG Arg	Leu	Arg	Ser 175	Cys	Asn	Thr	Ser \	/al 1 180	lle	
50	rnr	GIN	Ala	TGT Cys 185	Pro	Lys	Val	Thr	Phe 190	Glu :	Pro 1	le F	ro	
55	ATA Ile 195	CAT His	TAT Tyr	TGT Cys	Thr	CCG (Pro 2	GCT (Ala (GGT Gly	TTT (Ala :	ATT C Ile L 205	TG A	AG 6 .ys	521
60	TGT Cys	AAA Lys	GAT Asp 210	AAA Lys	AAG : Lys i	TTC / Phe /	Asn (GGA Gly 215	ACA (Thr (GGA (CCA T	ys L	AA 6 ys 20	60
-	AAT Asn	GTT Val	AGC Ser	ACA (GTA C	CAA 7	rgr A Cys 1	CA (His (GA A	TT A	AG C ys P	CA 6	99

	GTA Val	GTG Val 235	TCA Ser	ACT Thr	CAA Gln	CTG Leu	CTG Leu 240	TTA L u	AAT Asn	GGC Gly	AGC Ser	CTA Leu 245	GCA Ala	738
5	GAA Glu	GAA Glu	GAG Glu	GTA Val 250	ATA Ile	ATT Ile	AGA Arg	TCT Ser	GCC Ala 255	AAT Asn	TTC Phe	TCA Ser	AAC Asn	777
10	AAT Asn 260	GCT Ala	AAA Lys	ATC Ile	ATA Ile	ATA Ile 265	GTA Val	CAG Gln	TTG Leu	AAG Lys	GAA Glu 270	CCT Pro	GTA Val	816
15	GAA Glu	ATT Ile	AAT Asn 275	TGT Cys	ACA Thr	AGA Arg	CCC Pro	AGC Ser 280	AAC Asn	AAT Asn	ACA Thr	ATA Ile	AAA Lys 285	855
20	GGT Gly	ATA Ile	CAC His	ATA Ile	GGA Gly 290	CCA Pro	GGG Gly	AGA Arg	GCA Ala	TTT Phe 295	TAT Tyr	GCA Ala	ACA Thr	894
20	GGA Gly	GAC Asp 300	ATA Ile	CGA Arg	GGA Gly	GAT Asp	ATA Ile 305	AGA Arg	CAA Gln	GCA Ala	CAT His	TGT Cys 310	AAC Asn	933
25	ATT Ile	AGT Ser	GGA Gly	GCA Ala 315	AAA Lys	TGG Trp	AAT Asn	AAC Asn	ACT Thr 320	TTA Leu	AAG Lys	AAG Lys	GTA Val	972
30	GTT Val 325	ATA Ile	T y a	TTA Leu	AAA Lys	GAA Glu 330	Gln	TTT Phe	CCA Pro	AAT Asn	AAA Lys 335	ACA Thr	ATA Ile	1011
35	GTC Val	TTT Phe	AAC Asn 340	CAT His	TCC Ser	TCA Ser	GGA Gly	GGG Gly 345	GAC Asp	CCA Pro	GAA Glu	ATT Ile	GTA Val 350	1050
40	ATG Met	CAC His	AGT Ser	TTT Phe	AAT Asn 355	TGT Cys	CAA Gln	GGG Gly	GAA Glu	TTT Phe 360	TTC Phe	TAC Tyr	TGT Cys	1089
40	AAT Asn	ACA Thr 365	ACG Thr	AAG Lys	CTG Leu	TTT Phe	AAT Asn 370	AGT Ser	ACT Thr	TGG Trp	AAT Asn	GAT Asp 375	ACT Thr	1128
45	ACA Thr	GAG Glu	TCA Ser	AAT Asn 380	AAC Asn	AAT Asn	GAT Asp	AGT Ser	ACT Thr 385	ATT Ile	ACA Thr	CTC Leu	CCA Pro	1167
50	TGC Cys 390	Arg	Ile	AAA Lys	Gln	Ile	Ile	Asn	Met	Trp	Gln	GAA Glu	GTA Val	1206
55	GGA Gly	AAA Lys	GCA Ala 405	ATG Met	TAT Tyr	GCC Ala	CCT Pro	CCC Pro 410	ATC Ile	AGA Arg	GGA Gly	GAA Glu	ATT Ile 415	1245
. 0	AAA Lys	TGT Cys	TCA Ser	TCA Ser	AAT Asn 420	ATT Ile	ACA Thr	GGA Gly	CTA Leu	CTG Leu 425	TTA Leu	ACA Thr	AGA Arg	1284
60	GAT Asp	GGT Gly 430	Gly	ATT Ile	AAC Asn	Thr	AGC Ser 435	GAT Asp	GCC Ala	ACC Thr	GAG Glu	ACC Thr 440	TTC Phe	1323

	AGA Arg	CCC Pro	G GG	A GG y Gl 44	y Gl	A GA y As	T AT p Me	G AG	G GA G As 45	p As	T TC	Ģ AG P Ar	A AG	T 1:	362
5	GAA Glu 455	Leu	A TA'	T AA	A TA' s Ty	T AA r Ly 46	A GT s Va O	A GT l Va	G AA l Ly	A AI s Il	T GA e Gl 46	u Pr	A TT o Le	A 14	101
10	GGA Gly	GTA Val	A GC Ala 470	a Pro	C AC	C AA	G GC s Al	A AA a Ly: 47	s Ar	A AG g Ar	A GT g Va	G GT l Va	G CA 1 G1 48	n	40
15	AGA Arg	GAA Glu	AAA Lys	A AGI	A GC/ g Ala 489	a Va	A AC	A CT	A GG.	A GC y Al 49	a Me	G TT	C CT	T 14 u	79
20	GGG Gly	TTC Phe 495	Leu	GG#	A GC# / Ala	X TA	A AGG a Ser 500	TTC Phe 501	2	03					
	(2)	i) S	EQUE	NCE	CHAR	ACTE	ERIST	CICS:							
25	(x:	(B) T D) T	YPE:	Ami OGY:	no A Lir				NO: 4	1:				
30	Gly	Val	Pro	Val	Trp 5	Lys	: Glu	Ala	Thr	Th:	Thr	Leu	Phe	Cy:	s Al 1
	Ser	Asp	Ala	Lys	Ala 20	Tyr	Asp	Thr	Glu	Va] 25		Asn	Val	Tr	Al 3
35	Thr	His	Ala	Cys	Val 35	Pro	Thr	Asp	Pro	Asn 40		Gln	Glu	Met	: Va 49
	Leu	Glu	Asn	Val	Thr 50	Glu	Asp	Phe	Asn	Met 55		Lys	Asn	Asp	Met 60
40	Val	Glu	Gln	Met	His 65	Glu	Xaa	Ile	Ile	Ser 70		Trp	Asp	Gln	Ser 75
45					80		Leu			85					90
	Суз	Thr	Asn	Trp	Lys 95	Glu	Asn	Asp	Thr	Lys 100	Thr	Asn	Ser	Ser	Ser 105
50	Thr	Thr	Thr	Asn	Asn 110	Ser	Ser	Ala	Thr	Ala 115	Asn	Ser	Ser	Ser	Thr 120
	Thr '	Thr	Asn	Ser	Ser 125	Trp	Gly	Glu	Ile	Lys 130	Glu	Gly	Glu	Ile	Lys 135
55	Asn (Cys	Ser	Phe	Asn 140	Ile	Thr	Thr	Gly	Ile 145	Arg	Asp	Lys	Val	Lys 150
60	Lys (Glu	Tyr	Ala	Leu 155	Phe	Tyr	Ser	Leu	Asp 160	Val	Val	Pro	Ile	Glu 165
	Asn /				170					175					180
65	Ile 7	Chr (Gln	Ala	Cys 185	Pro	Lys	Val	Thr	Phe 190	Glu	Pro	Ile	Pro	Ile 195

	His	Tyr	Сув	Thr	Pro 200	Ala	Gly	Phe	Ala	Ile 205	Leu	Lys	Cys	Lys	Asp 210
5	Lys	Lys	Phe	Asn	Gly 215	Thr	Gly	Pro	Cys	Lys 220	Asn	Val	Ser	Thr	Val 225
	Gln	Cys	Thr	His	Gly 230	Ile	Lys	Pro	Val	Val 235	Ser	Thr	Gln	Leu	Leu 240
10	Leu	Asn	Gly	Ser	Leu 245	Ala	Glu	Glu	Glu	Val 250	Ile	Ile	Arg	Ser	Ala 255
	Asn	Phe	Ser	Asn	Asn 260	Ala	Lys	Ile	Ile	11e 265	Val	Gln	Leu	Lys	Glu 270
15	Pro	Val	Glu	Ile	Asn 275	Cys	Thr	Arg	Pro	Ser 280	Asn	Asn	Thr	Ile	Lys 285
20	Gly	Ile	His	Ile	Gly 290	Pro	Gly	Arg	Ala	Phe 295	Tyr	Ala	Thr	Gly	Asp 300
	Ile	Arg	Gly	Asp	11e 305	Arg	Gln	Ala	His	Cys 310	Asn	Ile	Ser	Gly	Ala 315
25	Lys	Trp	Asn	Asn	Thr 320	Leu	Lys	Lys	Val	Val 325	Ile	Lys	Leu	Lys	Glu 330
20	Gln	Phe	Pro	Asn	Lys 335	Thr	Ile	Val	Phe	Asn 340	His	Ser	Ser	Gly	Gly 345
30	Asp	Pro	Glu	Ile	Val 350	Met	His	Ser	Phe	Asn 355	Cys	Gln	Gly	Glu	Phe 360
35	Phe	Tyr	Cha	Asn	Thr 365	Thr	Lys	Leu	Phe	Asn 370	Ser	Thr	Trp	Asn	Asp 375
	Thr	Thr	Glu	Ser	Asn 380	Asn	Asn	Asp	Ser	Thr 385	Ile	Thr	Leu	Pro	Cys 390
40	Arg	Ile	Lys	Gln	Ile 395	Ile	Asn	Met	Trp	Gln 400	Glu	Val	Gly	Lys	Ala 405
45	Met	Tyr	Ala	Pro	Pro 410	Ile	Arg	Gly	Glu	Ile 415	Lys	Cys	Ser	Ser	Asn 420
45	Ile	Thr	Gly	Leu	Leu 425	Leu	Thr	Arg	Asp	Gly 430	Gly	Ile	Asn	Thr	Ser 435
50	Asp	Ala	Thr	Glu	Thr 440	Phe	Arg	Pro	Gly	Gly 445	Gly	Asp	Met	Arg	Asp 450
	Asn	Trp	Arg	Ser	Glu 455	Leu	Tyr	Lys	Tyr	Lys 460	Val	Val	Lys	Ile	Glu 465
55	Pro	Leu	Gly	Val	Ala 470	Pro	Thr	Lys	Ala	Lys 475	Arg	Arg	Val	Val	Gln 480
60	Arg	Glu	Lys	Arg	Ala 485	Val	Thr	Leu	Gly	Ala 490	Met	Phe	Leu	Gly	Phe 495
90	Leu	Gly	Ala	Xaa	Ser 500										
65	(2) 1	NFO	RMATI	ON F	OR S	EQ I	D NC	:5:							

65 (i) SEQUENCE CHARACTERISTICS:

5	(;		(B) 1 (C) 1 (D) 1	LENGT TYPE : STRAN TOPOL ENCE	Nuc IDEDI OGY:	cleid NESS: Lir	Ac: Sinear	id ngle		NO:	5:			
10	G	Va.	A CC1 l Pro	r GTA Val	TCC	AAA Lys	GAA Glu	A GCA	A ACC	C AC	C AC	r Le	A TT	T 37
10	TG1 Cys	C GC/	A TCA A Ser 15	A GAT Asp	GCT Ala	AAA Lys	GCA Ala	TAT Tyr	Asp	r AC	A GAG	G GT	A CAT l His	3
15	AAT Asr	GTT Val	TGG Trp	GCT Ala	ACA Thr 30	His	GCC	TGT Cys	CT#	CCC Pro	Thi	A GAG	C CCC	2 115
20	AAC Asn	CCF Pro 40	Gln	GAA Glu	GTA Val	GTA Val	TTG Leu 45	Glu	AAT Asn	CTA Val	ACA Thr	GAA Glu	a Asr	154
25	Phe	Asn	Met	TGG Trp 55	Lys	Asn	Asn	Met	Val 60	Glu	Gln	Met	. His	}
30	61u 65	Asp	lle	ATC Ile	Ser	70	Trp	Asp	Gln	Ser	Leu 75	Lys	Pro	
	Cys	Val	Lys 80		Thr	Pro	Leu	Cys 85	Vāl	Thr	Leu	Asn	Cys 90	
35	Thr	Asn	Leu	GAG Glu	Asn 95	Ala	Asn	Asn	Thr	Glu 100	Asn	Ala	Asn	
40	Asn	Thr 105	Asn	AAT Asn	Tyr	Thr	Leu 110	Gly	Met	Glu	Arg	Gly 115	Glu	
45	Ile	Lys	Asn	TGC Cys 120	Ser	Phe	Asn	Ile	Thr 125	Thr	Ser	Leu	Arg	
50	Asp 130	Lys	Val	AAA Lys	Lys	Glu 135	Tyr	Ala	Leu	Phe	Tyr 140	Lys	Leu	
	GAT Asp	GTA Val	GTA Val 145	CAA Gln	ATA Ile	GAT Asp	AAT Asn	AGT Ser 150	ACC Thr	AAC Asn	TAT Tyr	AGG Arg	CTG Leu 155	466
55	ATA Ile	AGT Ser	TGT Cys	AAT Asn	ACC Thr 160	TCA Ser	GTC Val	ATT Ile	ACA Thr	CAG Gln 165	GCC Ala	TGT Cys	CCA Pro	505
60	AAG Lys	GTA Val 170	TCC Ser	TTT Phe	GAG Glu	Leu	ATT Ile 175	CCC Pro	ATA Ile	CAT His	TAT Tyr	TGT Cys 180	GCC Ala	544
65	CCG Pro	GCT Ala	Gly	TTT Phe	GCG Ala	ATT Ile	CTA Leu	Lys	TGT Cys 190	AAA Lys	GAT Asp	AAG Lys	AAG Lys	583

	TTC Phe 195	TAA Asn	GGA Gly	ACA Thr	GGA Gly	CCA Pro 200	TGT Cys	AAA Lys	AAT Asn	GTC Val	AGC Ser 205	ACA Thr	GTA Val	622
5	CAA Gln	TGT Cys	ACA Thr 210	CAT His	GGA Gly	ATT Ile	AGA Arg	CCA Pro 215	GTA Val	GTA Val	TCA Ser	ACT Thr	CAA Gln 220	661
10	CTA Leu	CTG Leu	TTA Leu	AAT Asn	GGC Gly 225	AGT Ser	CTA Leu	GCA Ala	GAA Glu	GAA Glu 230	GAG Glu	ATA Ile	GTA Val	700
15	ATT Ile	AGA Arg 235	TCT Ser	GAA Glu	AAT Asn	ATC Ile	ACA Thr 240	GAC Asp	AAT Asn	GCT Ala	AAA Lys	ACC Thr 245	ATA Ile	739
20	ATA Ile	GTG Val	CAG Gln	CTA Leu 250	AAT Asn	GAA Glu	TCT Ser	ATA Ile	GTG Val 255	ATT Ile	AAT Asn	TGT Cys	ACA Thr	778
20	AGA Arg 260	CCC Pro	AAT Asn	AAC Asn	AAC Asn	ACA Thr 265	AGA Arg	AAA Lys	AGT Ser	ATA Ile	AAT Asn 270	ATA Ile	GGA Gly	817
25	CCA Pro	GGG Gly	AGA Arg 275	GCA Ala	TTC Phe	TAT Tyr	ACA Thr	ACA Thr 280	GGA Gly	Aab	ATA Ile	ATA Ile	GGA Gly 285	856
30	GAT Asp	ATA Ile	AGA Arg	CAA Gln	GCA Ala 290	CAT His	TGT Cys	AAC Asn	CTT Leu	AGT Ser 295	AAA Lys	ACA Thr	CAA Gln	895
35	TGG Trp	GAA Glu 300	AAA Lys	ACG Thr	TTA Leu	AGA Arg	CAG Gln 305	ATA Ile	GCT Ala	ATA Ile	AAA Lys	TTA Leu 310	GAA Glu	934
4.0	GAA Glu	AAA Lys	TTT Phe	AAG Lys 315	AAT Asn	AAA Lys	ACA Thr	ATA Ile	GCC Ala 320	TTT Phe	AAT Asn	AAA Lys	TCC Ser	973
40	TCA Ser 325	GGA Gly	GGG Gly	GAC Asp	CCA Pro	GAA Glu 330	ATT Ile	GTA Val	ATG Met	CAC His	AGT Ser 335	TTT Phe	AAT Asn	1012
45	TGT Cys	GGA Gly	GGG Gly 340	GAA Glu	TTT Phe	TTC Phe	TAC Tyr	TGT Cys 345	AAT Asn	ACA Thr	ACA Thr	AAA Lys	CTG Leu 350	1051
50	TTT Phe	AAT Asn	AGT Ser	ACC Thr	TGG Trp 355	AAT Asn	TTA Leu	ACA Thr	CAA Gln	CCG Pro 360	TTT Phe	AGT Ser	AAT Asn	1090
55	ACC Thr	GGG Gly 365	AAT Asn	CGT Arg	ACT Thr	GAA Glu	GAG Glu 370	TTA Leu	AAT Asn	ATT Ile	ACA Thr	CTC Leu 375	CCA Pro	1129
	TGC Cys	AGA Arg	ATA Ile	AAA Lys 380	CAA Gln	ATC Ile	ATA Ile	AAC Asn	TTG Leu 385	TGG Trp	CAG Gln	GAA Glu	GTA Val	1168
60	GGC Gly 390	AAA Lys	GCA Ala	ATG Met	TAT Tyr	GCC Ala 395	CCT Pro	CCC Pro	ATC Ile	AGA Arg	GGA Gly 400	CAA Gln	ATT Ile	1207

	AGA Arg	TGT Cys	TCA Ser 405	TCA Ser	AAT Asn	ATT	ACA Thr	GGG Gly 410	Leu	Leu	TTA Leu	ACA Thr	A AGA Arc 419)	46
5	GAT Asp	GGT Gly	GGA Gly	AGT Ser	AAC Asn 420	ACC Thr	GGT Gly	Asp	AAC Asn	AGG Arg 425	ACT Thr	GAG Glu	ACC Thr	128	85
10											AAT Asn		Arg		24
15										Arg	ATT				53
20	TTA Leu 455	GGA Gly	GTA Val	GCA Ala	CCC Pro	ACC Thr 460	CAG Gln	GCA Ala	AAG Lys	AGA Arg	AGA Arg 465	GTG Val	GTG Val	140)2
20	CAA Gln	AGA Arg	GAA Glu 470	AAA Lys	AGA Arg	GCA Ala	GTG Val	GGG Gly 475	ATA Ile	GGA Gly	GCT Ala	ATG Met	TTC Phe 480		1
25			TTC Phe				AA :	1461							
30	(2) INFORMATION FOR SEQ ID NO:6:														
35) SE	QUEN	CE D	ESCR	IPTI	ON:								
	1				5					10			-		Ser 15
4 0					20					25			_		Thr 30
	His .	Ala	Cys	Val	Pro 35	Thr	Asp	Pro	Asn	Pro 40	Gln	Glu	Val	Val	Leu 45
45	Glu	Asn	Val	Thr	Glu 50	Asn	Phe	Asn	Met	Trp 55	Lys	Asn	Asn	Met	Val 60
50	Glu	Gln .	Met	His	Glu 65	Asp	Ile	Ile	Ser	Leu 70	Trp	Asp	Gln	Ser	Leu 75
,	Lys	Pro	Cys	Val :	Lys 80	Leu	Thr	Pro	Leu	Cys 85	Val	Thr	Leu	Asn	Cys 90
55	Thr A	Asn :	Leu (Glu .	Asn 95	Ala	Asn	Asn	Thr	Glu 100	Asn	Ala	Asn	Asn	Thr 105
	Asn i	Asn '	Tyr '		Leu 110	Gly .	Met	Glu	Arg	Gly 115	Glu	Ile	Lys	Asn	Cys 120
50	Ser 1	Phe i	Asn :		Thr '	Thr .	Ser	Leu	Arg	Asp 130	l.ys	Val	Lys	Lys	Glu 135
55	Tyr i	Ala 1	Leu 1		Tyr :	Lys :	Leu .	qeA		Val 145	Gln	Ile	Asp	Asn	Ser 150

	Thr	Asn	Tyr	Arg	Leu 155	Ile	Ser	Cys	Asn	Thr 160	Ser	Val	Ile	Thr	Gln 165
5	Ala	Сув	Pro	Lys	Val 170	Ser	Phe	Glu	Leu	Ile 175	Pro	Ile	His	Tyr	Cys 180
	Ala	Pro	Ala	Gly	Phe 185	Ala	Ile	Leu	Lys	Cys 190	Lys	Asp	Lys	Lys	Phe 195
10	Asn	Gly	Thr	Gly	Pro 200	Суз	Lys	Asn	Val	Ser 205	Thr	Val	Gln	Cys	Thr 210
	His	Gly	Ile	Arg	Pro 215	Val	Val	Ser	Thr	Gln 220	Leu	Leu	Leu	Asn	Gly 225
15	Ser	Leu	Ala	Glu	Glu 230	Glu	Ile	Val	Ile	Arg 235	Ser	Glu	Asn	Ile	Thr 240
20	Asp	Asn	Ala	Lys	Thr 245	Ile	Ile	Val	Gln	Leu 250	Asn	Glu	Ser	lle	Val 255
	Ile	Asn	Cys	Thr	Arg 260	Pro	Asn	Asn	Asn	Thr 265	Arg	Lys	Ser	Ile	Asn 270
25	Ile	Gly	Pro	Gly	Arg 275	Ala	Phe	Tyr	Thr	Thr 280	Gly	Asp	Ile	Ile	Gly 285
7.0	Asp	Ile	Arg	Gln	Ala 290	His	Cys	Asn	Leu	Ser 295	Lys	Thr	Gln	Trp	Glu 300
30	Lys	Thr	Leu	Arg	Gln 305	Ile	Ala	Ile	Lys	Leu 310	Glu	Glu	Lys	Phe	Lys 315
35	Asn	Lys	Thr	Ile	Ala 320	Phe	Asn	Lys	Ser	Ser 325	Gly	Gly	Asp	Pro	Glu 330
	Ile	Val	Met	His	Ser 335	Phe	Asn	Cys	Gly	Gly 340	Glu	Phe	Phe	Tyr	Cys 345
40	Asn	Thr	Thr	Lys	Leu 350	Phe	Asn	Ser	Thr	Trp 355	Asn	Leu	Thr	Gln	Pro 360
4. 5	Phe	Ser	Asn	Thr	Gly 365	Asn	Arg	Thr	Glu	Glu 370	Leu	Asn	Ile	Thr	Leu 375
45	Pro	Суз	Arg	Ile	Lys 380	Gln	Ile	Ile	Asn	Leu 385	Trp	Gln	Glu	Val	Gly 390
50	Lys	Ala	Met	Tyr	Ala 395	Pro	Pro	Ile	Arg	Gly 400	Gln	Ile	Arg	Сув	Ser 405
	Ser	Asn	Ile	Thr	Gly 410	Leu	Leu	Leu	Thr	Arg 415	Asp	Gly	Gly	Ser	Asn 420
55	Thr	Gly	Asp	Asn	Arg 425	Thr	Glu	Thr	Phe	Arg 430	Pro	Gly	Gly	Gly	Asp 435
	Met	Arg	Asp	Asn	Trp 440	Arg	Ser	Glu	Leu	Tyr 445	Lys	Tyr	Lys	Val	Val 450
60	Arg	Ile	Glu	Pro	Leu 455	Gly	Val	Ala	Pro	Thr 460	Gln	Ala	Lys	Arg	Arg 465
65	Val	Val	Gln	Arg	Glu 470	Lys	Arg	Ala	Val	Gly 475	Ile	Gly	Ala	Met	Phe 480

Leu Gly Phe Leu Gly Asp 485 486

5	(2) INFORMATION FOR SEQ ID NO:7: (i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 1474 base pairs (B) TYPE: Nucleic Acid (C) STRANDEDNESS: Single													
10	(xi)	(D)	TOPO	LOGY	: Li	near	-		NO:	7:			
15	G	va	A CC	T GT(G TG	b ra	A GA s Gl	A GC.	A AC a Th	C AC r Th	C AC r Th	T CT/ r Let	A TT	r 37
20	TG: Cy:	T GC	A TC	r Ası	r GC	T AA a Lys	A GC	A TA' a Ty: 20	r As	T AC	A GAG	G GT#	A CA1	;
20	AA: Asi	r Gr	T TGC	G GCT	ACA Tha	c HT8	Γ GC0	TG1 Cys	CT/	A CCC l Pro	Thi	A GAC	CCC Pro	115
25	AAC Asr	C CCA Pro 40	GIL	GAA Glu	GT#	GTA Val	TTC Leu	: Glu	AAT Asr	GT#	ACA Thr	GAA Glu 50	Asn	154
30	TT1 Phe	AAC Asn	ATG Met	TGG Trp 55	Lys	AAT ASn	`AAC Asn	ATG Met	GTA Val	Glu	CAG Gln	ATG Met	CAT His	193
35	GAG Glu 65	Asp	ATA Ile	ATC Ile	AGT Ser	TTA Leu 70	Trp	GAT Asp	CAA Gln	AGT Ser	CTA Leu 75	AAG Lys	CCA Pro	232
40	TGT Cys	GTA Val	AAA Lys 80	Leu	ACC Thr	CCA Pro	CTC Leu	TGT Cys 85	GTT Val	ACT Thr	TTA Leu	AAT Asn	TGC Cys 90	271
	ACT Thr	AAT Asn	TTG Leu	GAG Glu	AAT Asn 95	GCT Ala	AAT Asn	AAT Asn	ACC Thr	GAG Glu 100	AAT Asn	GCT Ala	AAT Asn	310
45	AAT Asn	ACC Thr 105	AAT Asn	AAT Asn	TAT Tyr	ACC Thr	TTG Leu 110	GGG Gly	ATG Met	GAG Glu	AGA Arg	GGT Gly 115	GAA Glu	349
50	AGA Arg	AAA Lys	AAC Asn	TGC Cys 120	TCT Ser	TTC Phe	AAT Asn	ATC Ile	ACC Thr 125	ACA Thr	AGC Ser	TTA Leu	AGA Arg	388
55	GAT Asp 130	AAG Lys	GGG Gly	AAA Lys	AAA Lys	GAA Glu 135	TAT Tyr	GCA Ala	TTG Leu	TTT Phe	TAT Tyr 140	AAA Lys	CTT Leu	427
60	GAT Asp	GTA Val	GTA Val 145	CAA Gln	ATA Ile	GAT Asp	AAT Asn	AGT Ser 150	ACC Thr	AAC Asn	TAT Tyr	AGG Arg	CTG Leu 155	466
	ATA Ile	AGT Ser	TGT Cys	ASN	ACC Thr 160	TCA Ser	GTC Val	ATT Ile	Thr	CAG Gln 165	GCC Ala	TGT (Cys)	CCA ! Pro	505

	AAG Lys	GTA Val 170	TCC Ser	TTT Phe	GAG Glu	CCA Pro	ATT Ile 175	CCC Pro	ATA Ile	CAT His	TAT Tyr	TGT Cys 180	GCC Ala	544
5	CCG Pro	GCT Ala	GGT Gly	TTT Phe 185	GCG Ala	ATT Ile	CTA Leu	AAG Lys	TGT Cys 190	AAA Lys	GAT Asp	AAG Lys	AAG Lys	583
10	TTC Phe 195	AAT Asn	GGA Gly	ACA Thr	GGA Gly	CCA Pro 200	TGT Cys	AAA Lys	AAT Asn	GTC Val	AGG Arg 205	ACA Thr	GTA Val	622
15	CAA Gln	TGT Cys	ACA Thr 210	CAT His	GGA Gly	ATT Ile	AGA Arg	CCA Pro 215	GTA Val	GTA Val	TCA Ser	ACT Thr	CAA Gln 220	661
	CTA Leu	CTG Leu	TTA Leu	AAT Asn	GGC Gly 225	AGT Ser	CTA Leu	GCA Ala	GAA Glu	GAA Glu 230	GAG Glu	ATA Ile	GTA Val	700
20	ATT Ile	AGA Arg 235	TCT Ser	GAA Glu	AAT Asn	ATC Ile	ACA Thr 240	GAC Asp	AAT Asn	GCT Ala	AAA Lys	ACC Thr 245	ATA Ile	739
25	ATA Ile	GTG Val	CAG Gln	CTA Leu 250	AAT Asn	GAA Glu	TCT Ser	ATA Ile	GTG Val 255	ATT Ile	AAT Asn	TGT Cys	ACA Thr	778
30	AGA Arg 260	CCC Pro	AAT Asn	AAC Asn	AAC Asn	ACA Thr 265	AGA Arg	AAA Lys	AGT Ser	ATA Ile	AAT Asn 270	ATA Ile	GGA Gly	817
35	CCA Pro	GGG Gly	AGA Arg 275	GCA Ala	TTC Phe	TAT Tyr	ACA Thr	ACA Thr 280	GGA Gly	GAC Asp	ATA Ile	ATA Ile	GGA Gly 285	856
	GAT Asp	ATA Ile	AGA Arg	CAA Gln	GCA Ala 290	CAT His	TGT Cys	AAC Asn	CTT Leu	AGT Ser 295	AAA Lys	ACA Thr	CAA Gln	895
40	TGG Trp	GAA Glu 300	AAA Lys	ACG Thr	TTA Leu	AGA Arg	CAG Gln 305	ATA Ile	GCT Ala	ATA Ile	AAA Lys	TTA Leu 310	GAA Glu	934
45	GAA Glu	AAA Lys	TTT Phe	AAG Lys 315	AAT Asn	AAA Lys	ACA Thr	ATA Ile	GCC Ala 320	TTT Phe	AAT Asn	AAA Lys	TCC Ser	973
50								GTA Val						1012
55	TGT Cys	GGA Gly	GGG Gly 340	GGA Gly	TTT Phe	TTC Phe	TAC Tyr	TGT Cys 345	AGT Ser	ACG Thr	AGA Arg	AAA Lys	CTG Leu 350	1051
	TTT Phe	AAT Asn	AGT Ser	ACC Thr	TGG Trp 355	AAT Asn	TTA Leu	ACA Thr	CAA Gln	CCG Pro 360	TTT Phe	AGT Ser	AAT Asn	1090
60	ACC Thr	GGG Gly 365	GAT Asp	CGT Arg	ACT Thr	GAA Glu	GAG Glu 370	TTA Leu	AAT Asn	ATT Ile	ACA Thr	CTC Leu 375	CCA Pro	1129

	TGC Cys	AGA Arg	ATA Ile	AAA Lya 380	Gln	ATC Ile	ATA E Ile	A AAC Pass	7TC Leu 389	TGC Trp	CAC Gli	G GA n Gl	A GT u Va	A 11	168
5	GGC Gly 390	AAA Lys	GCA Ala	ATG Met	TAT	GCC Ala 395	Pro	CCC Pro	ATC Ile	AGA Arg	GG/ Gly 400	/ Gl:	A AT	T 12 e	207
10	AGA Arg	TGT Cys	TCA Ser 405	TCA Ser	AAT Asn	ATT	ACA Thr	GGG Gly 410	Leu	CTA Leu	Leu	A AGO	AG Arc 41	3	46
15	GAT Asp	GGT Gly	GGA Gly	AGT Ser	AAC Asn 420	ACC Thr	AGT Ser	GAC Asp	AAC Asn	CAG Gln 425	ACT Thr	GAC	ACC Thi	12	85
20	TTT Phe	AGA Arg 430	CCT Pro	GGG Gly	GGA Gly	GGA Gly	GAT Asp 435	ATG Met	AGG Arg	GAC Asp	AAG Lys	TGG Trp 440	Arç	\ 13 J	24
	AGT Ser	GAA Glu	TTA Leu	TAT Tyr 445	AAA Lys	TAT Tyr	AAA Lys	GTA Val	GTA Val 450	AGA Arg	ATT	GAA Glu	CCA Pro	130	63
25	TTA Leu 455	GGA Gly	GTA Val	GCA Ala	CCC Pro	ACC Thr 460	CAG Gln	GCA Ala	AAG Lys	AGA Arg	AGA Arg 465	GTG Val	CTG Val	140	02
30	CAA Gln	AGA Arg	GAA Glu 470	AAA Lys	ÀGA Arg	GCA Ala	GTG Val	GGG Gly 475	ATA Ile	GGA Gly	GCT Ala	ATG Met	TTC Phe 480	144	11
35	CTT .	AGG Arg	TTC Phe	Leu	GGA Gly 485	GAT Asp	AAA Lys	GCT Ala	TCT Ser	AGA Arg 490	Val	1474	4		
40	(2) INFORMATION FOR SEQ ID NO:8: (i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 491 amino acids (B) TYPE: Amino Acid (D) TOPOLOGY: Linear (xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:														
45	Val I				5					10					15
50	Asp A				20					25					30
50	His A				35					40					45
55	Glu A				50					55					60
	Glu G				65					70					75
60	Lys P				80					85					90
	Thr A	sn L	eu G	lu A	sn A 95	la A	Asn A	Asn I		100 100	sn A	Ala i	Asn .	Asn	Thr 105

	Asn	Asn	Tyr	Thr	Leu 110	Gly	Met	Glu	Arg	Gly 115	Glu	Arg	Lys	Asn	Cys 120
5	Ser	Phe	Asn	Ile	Thr 125	Thr	Ser	Leu	Arg	Asp 130	Lys	Gly	Lys	Lys	Glu 135
	Tyr	Ala	Leu	Phe	Tyr 140	Lys	Leu	Asp	Val	Val 145	Gln	Ile	Asp	Asn	Ser 150
10	Thr	Asn	Tyr	Arg	Leu 155	Ile	Ser	Cys	Asn	Thr 160	Ser	Val	Ile	Thr	Gln 165
	Ala	Cys	Pro	Lys	Val 170	Ser	Phe	Glu	Pro	Ile 175	Pro	Ile	His	Tyr	Cys 180
15	Ala	Pro	Ala	Gly	Phe 185	Ala	Ile	Leu	Lys	Cys 190	Lys	Asp	Lys	Lys	Phe 195
20	Asn	Gly	Thr	Gly	Pro 200	Cys	Lys	Asn	Val	Arg 205	Thr	Val	Gln	Cys	Thr 210
	His	Gly	Ile	Arg	Pro 215	Val	Val	Ser	Thr	Gln 220	Leu	Leu	Leu	Asn	Gly 225
25	Ser	Leu	Ala	Glu	Glu 230	Glu	Ile	Val	Ile	Arg 235	Ser	Glu	Asn	Ile	Thr 240
20	Asp	Asn	Ala	Lys	Thr 245	Ile	Ile	Val	Gln	Leu 250	Asn	Glu	Ser	Ile	Val 255
30	Ile	Asn	Cys	Thr	Arg 260	Pro	Asn	Asn	Asn	Thr 265	Arg	Lys	Ser	Ile	Asn 270
35	Ile	Gly	Pro	Gly	Arg 275	Ala	Phe	Tyr	Thr	Thr 280	Gly	Asp	Ile	Ile	Gly 285
	Asp	Ile	Arg	Gln	Ala 290	His	Cys	Asn	Leu	Ser 295	Lys	Thr	Gln	Trp	Glu 300
40	Lys	Thr	Leu	Arg	Gln 305	Ile	Ala	Ile	Lys	Leu 310	Glu	Glu	Lys	Phe	Lys 315
4.5	Asn	Lys	Thr	Ile	Ala 320	Phe	Asn	Lys	Ser	Ser 325	Gly	Gly	Asp	Pro	Glu 330
45	Ile	Val	Met	His	Ser 335	Phe	Asn	Cys	Gly	Gly 340	Gly	Phe	Phe	Tyr	Cys 345
50	Ser	Thr	Arg	Lys	Leu 350	Phe	Asn	Ser	Thr	Trp 355	Asn	Leu	Thr	Gln	Pro 360
	Phe	Ser	Asn	Thr	Gly 365	Asp	Arg	Thr	Glu	Glu 370	Leu	Asn	Ile	Thr	Leu 375
55	Pro	Cys	Arg	Ile	Lys 380	Gln	Ile	Ile	Asn	Leu 385	Trp	Gln	Glu	Val	Gly 390
60	Lys	Ala	Met	Tyr	Ala 395	Pro	Pro	Ile	Arg	Gly 400	Gln	Ile	Arg	Cys	Ser 405
60	Ser	Asn	Ile	Thr	Gly 410	Leu	Leu	Leu	Arg	Arg 415	Asp	Gly	Gly	Ser	Asn 420
65	Thr	Ser	Asp	Asn	Gln 425	Thr	Glu	Thr	Phe	Arg 430	Pro	Gly	Gly	Gly	Asp 435

	Met	Arg A	sp Ly	s Tr 44	p Ar	g Se	r Gl	u Le	u Ty 44		s Ty	r Ly	s Val	Val 450
5	Arg	Ile G	lu Pr	o Le 45	u Gl	y Va	l Al	a Pr	o Th 46	r Gl	n Al	a Ly	s Arg	Arg 465
	Val	Val G	ln Ar	g G1:	u Ly: O	s Ar	g Al	a Va	1 G1 47	y Il.	e Gl	y Al	a Met	Phe 480
10	Leu	Arg P	ne Le	48		p Ly:	s Al	a Se		y Va:				
15	(i	(B) (C)	JENCE LENG TYPE STRAI TOPOI	CHAI TH: Nuc NDEDI LOGY:	RACTE 1512 cleic NESS: Lin	ERIST base Aci Sir ear	rics pa id ngle	: irs	NO: 9) <u>:</u>				
20	,	CTC GA Leu Gl	G GT	CCI	GTA	TGC Trp	. AA	A GAA	GCA	ACT	Thr	ACT Thr	36	
25 .	Leu	1	s Ala 5	Ser	. Asb	Ala	Lys 20	Ala	Tyr	Asn	Thr	Glu 25		
30	AAA (Lys !	CAT AA His As	T GTT n Val	TGG Trp 30	Ala	ACA Thr	CAC	GCC Ala	TGT Cys 35	Val	CCC Pro	ACA Thr	114	
35	GAT (Asp I	CCC AA Pro As 40	C CCA n Pro	CAA Gln	GAA Glu	GTA Val 45	GTA Val	TTG Leu	GGA Gly	AAT Asn	GTG Val 50	ACA Thr	153	
40	Glu A	AAT TT Asn Ph	e Asn 55	Met	Trp	Lys	Asn	Asn 60	Met	Val	Glu	Gln		
	ATG C Met H 65	CAT GA	A GAT u Asp	ATA Ile	ATC Ile 70	AGT Ser	TTA Leu	TGG Trp	GAT Asp	CAA Gln 75	AGT Ser	CTA Leu	231	
45	AAG C Lys P	CA TG	s Val	AAA Lys	TTA Leu	ACC Thr	CCA Pro 85	CTC Leu	TCT Cys	CTT Val	ACT Thr	TTA Leu 90	270	
50	AAT T Asn C	GC AC	GAT Asp	GAT Asp 95	TTA Leu	GGG Gly	AAT Asn	GCT Ala	ACT Thr 100	AAT Asn	ACC Thr	AAT Asn	309	
55	Ser S	GT GCC er Ala 05	Thr	Thr	Asn	Ser 110	Ser	Ser	Trp	Glu	Glu 115	Met		
60	AAG G	GG GAA ly Glu	ATG Met 120	AAA Lys	AGA Arg	TGC Cys	TCT Ser	TTC Phe 125	AAT Asn	ATC Ile	ACC Thr	ACA Thr	387	•
	AGC ASSET I	TA AGA le Arg	GAT Asp	AAG Lys	ATT Ile 135	AAG Lys	AAA Lys	GAA Glu	His	GCA Ala 140	CTT Leu	TTC Phe	426	

	TAT Tyr	AGA Arg	CTT Leu 145	GAT Asp	GTA Val	GTA Val	CCA Pro	ATA Ile 150	GAT Asp	AAT Asn	GAT Asp	AAT Asn	ACC Thr 155	465
5	ACA Thr	TAT Tyr	AGG Arg	TTG Leu	ATA Ile 160	AAT Asn	TGT Cys	AAT Asn	ACC Thr	TCA Ser 165	GTC Val	ATT Ile	ACA Thr	504
10	CAG Gln	GCC Ala 170	TGT Cys	CCA Pro	AAG Lys	GTA Val	TCA Ser 175	TTT Phe	GAG Glu	CCA Pro	ATT Ile	CCC Pro 180	ATA Ile	543
15	CAT His	TTT Phe	TGT Cys	GCC Ala 185	CCG Pro	GCT Ala	GGT Gly	TTT Phe	GCG Ala 190	ATT Ile	CTA Leu	AAG Lys	TGT Cys	582
20	AAT Asn 195	AAT Asn	AAG Lys	ACG Thr	TTC Phe	GAG Glu 200	GGA Gly	AAA Lys	GGA Gly	CCA Pro	TCT Cys 205	AAA Lys	AAT Asn	621
20	GTC Val	AGT Ser	ACA Thr 210	GTA Val	CAA Gln	TGC Cys	ACA Thr	CAT His 215	GGA Gly	ATT Ile	AGG Arg	CCA Pro	GTA Val 220	660
25	GTG Val	TCA Ser	ACT Thr	CAA Gln	CTG Leu 225	CTG Leu	TTA Leu	AAT Asn	GGC Gly	AGT Ser 230	CTA Leu	GCA Ala	GAA Glu	699
30	GAA Glu	GAG Glu 235	GTA Val	ATA Ile	ATT Ile	AGA Arg	TCT Ser 240	GAC Asp	AAT Asn	ATC Ile	ACA Thr	GAC Asp 245	AAT Asn	738
35	ACT Thr	AAA Lys	ACC Thr	ATT Ile 250	ATA Ile	GTA Val	CAG Gln	CTA Leu	AAC Asn 255	GAA Glu	TCT Ser	GTA Val	GTA Val	777
•	ATT Ile 260	AAT Asn	TGT Cys	ACA Thr	AGA Arg	CCC Pro 265	AAC Asn	AAC Asn	AAT Asn	ACA Thr	AGA Arg 270	AAA Lys	AGT Ser	816
40	ATA Ile	CAT His	ATA Ile 275	GGA Gly	CCA Pro	GGG Gly	AGT Ser	GCA Ala 280	TTT Phe	TTT Phe	GCA Ala	ACA Thr	GGA Gly 285	855
45	GAA Glu	ATA Ile	ATA Ile	GGA Gly	GAT Asp 290	ATA Ile	AGA Arg	CAA Gln	GCA Ala	CAC His 295	TGT Cys	AAC Asn	CTT Leu	894
50	AGT Ser	AGA Arg 300	ACA Thr	CAA Gln	TGG Trp	AAT Asn	AAC Asn 305	ACT Thr	TTA Leu	GGA Gly	AAG Lys	ATA Ile 310	GTC Val	933
55	ATA Ile	AAA Lys	TTA Leu	AGA Arg 315	GAA Glu	CAA Gln	TTT Phe	AGA Arg	AAA Lys 320	CAA Gln	TTT Phe	GGA Gly	GAA Glu	972
40	AAA Lys 325	ACA Thr	ATA Ile	GTC Val	TTT Phe	AAT Asn 330	CGA Arg	TCC Ser	TCA Ser	GGA Gly	GGG Gly 335	GAC Asp	CCG Pro	1011
60	GAA Glu	ATT Ile	GCA Ala 340	ATG Met	CAC His	AGT Ser	TTT Phe	AAT Asn 345	TGT Cys	GGA Gly	GGG Gly	GAA Glu	TTT Phe 350	1050

	TTC Phe	TAC Tyr	TGT Cys	AAC Asn	ACA Thr 355	ACA Thr	GCA Ala	CTG Leu	TTT Phe	AAT Asn 360	Ser	Thr	TG(3 10 P	89
5	TAA neA	GTT Val 365	Thr	AAA Lys	GGG Gly	TTG Leu	AAT Asn 370	Asn	ACT Thr	GAA Glu	GGA Gly	AAT Asn 375	Sei	2 11	28
10	ACA Thr	GGA Gly	GAT Asp	GAA Glu 380	TAA neA	ATC Ile	ATA Ile	CTC Leu	CCA Pro 385	Cys	AGA Arg	ATA Ile	AA? Lys	A 110	67
15	CAA Gln 390	Ile	ATA Ile	AAC Asn	ATG Met	TGG Trp 395	CAG Gln	GAA Glu	GTA Val	GGA Gly	AAA Lys 400	Ala	ATO	120	06
20	TAT Tyr	GCC Ala	CCT Pro 405	CCC Pro	ATC Ile	AGT Ser	GGA Gly	CAA Gln 410	ATT Ile	AGA Arg	TGT Cys	TCA Ser	TCA Ser 415		4 5
20							CTA Leu								34
25	AAG Lys	AAC Asn 430	GAG Glu	AGC Ser	ATC Ile	ACC Thr	ACC Thr 435	GAG Glu	GTC Val	TTC Phe	AGA Arg	CCT Pro 440	GGA Gly	132	23
30	GGA Gly	GGA Gly	GAT Asp	ATG Met 445	AGG Arg	GAC Asp	AAT Asn	TGG Trp	AGA Arg 450	AGT Ser	GAA Glu	TTA Leu	TAT Tyr	136	2
35	AAA Lys 455	TAT Tyr	AAA Lys	GTA Val	GTA Val	AAA Lys 460	ATT Ile	GAA Glu	CCA Pro	TTA Leu	GGA Gly 465	GTA Val	GCG Ala	140	1
40							AGA Arg							144	0
40							GGA Gly							147	9
45							TAG Xaa 500					1512	?		
50	(2) 1 (i	i) SE (A (E	QUEN	CE C NGTH PE:	HARA : 50 Amin	CTER 4 am o Ac	ISTI ino id	cs:	s						
55	(×i	•	•				ON:	SEQ	ID N	0:10	:				
	1				5		Lys			10					15
60	Ala	Ser	Asp	Ala	Lys 20	Ala	Tyr	Asn '	Thr	G1u 25	Lys	His	Asn	Val	Trp 30
	Ala	Thr	His	Ala	Cys '	Val	Pro '	Thr	Asp	Pro	Asn	Pro	Gln	Glu	Val

	Val	Leu	Gly	Asn	Val 50	Thr	Glu	Asn	Phe	Asn 55	Met	Trp	Lys	Asn	Asn 60
5	Met	Val	Glu	Gln	Met 65	His	Glu	Asp	Ile	11e 70	Ser	Leu	Trp	Asp	Gln 75
	Ser	Leu	Lys	Pro	Cys 80	Val	Lys	Leu	Thr	Pro 85	Leu	Cys	Val	Thr	Leu 90
10	Asn	Cys	Thr	Asp	Asp 95	Leu	Gly	Asn	Ala	Thr 100	Asn	Thr	Asn	Ser	Ser 105
15	Ala	Thr	Thr	Asn	Ser 110	Ser	Ser	Trp	Glu	Glu 115	Met	Lys	Gly	Glu	Met 120
15	Lys	Arg	Cys	Ser	Phe 125	Asn	Ile	Thr	Thr	Ser 130	Ile	Arg	Asp	Lys	Ile 135
20	Lys	Lys	Glu	His	Ala 140	Leu	Phe	Tyr	Arg	Leu 145	Asp	Val	Val	Pro	Ile 150
	Asp	Asn	Asp	Asn	Thr 155	Thr	Tyr	Arg	Leu	11e 160	Asn	Cys	Asn	Thr	Ser 165
25	Val	Ile	Thr	Gln	Ala 170	Cys	Pro	Lys	Val	Ser 175	Phe	Glu	Pro	Ile	Pro 180
20	Ile	His	Phe	Cys	Ala 185	Pro	Ala	Gly	Phe	Ala 190	Ile	Leu	Lys	Cys	Asn 195
30	Asn	Lys	Thr	Phe	Glu 200	Gly	Lys	Gly	Pro	Cys 205	Lys	Asn	Val	Ser	Thr 210
35	Val	Gln	Cys	Thr	His 215	Gly	Ile	Arg	Pro	Val 220	Val	Ser	Thr	Gln	Leu 225
	Leu	Leu	Asn	Gly	Ser 230	Leu	Ala	Glu	Glu	Glu 235	Val	Ile	Ile	Arg	Ser 240
40	Asp	Asn	Ile	Thr	Asp 245	Asn	Thr	Lys	Thr	11e 250	Ile	Val	Gln	Leu	Asn 255
45	Glu	Ser	Val	Val	Ile 260	Asn	Cys	Thr	Arg	Pro 265	Asn	Asn	Asn	Thr	Arg 270
45	Lys	Ser	Ile	His	Ile 275	Gly	Pro	Gly	Ser	Ala 280	Phe	Phe	Ala	Thr	Gly 285
50	Glu	Ile	Ile	Gly	Asp 290	Ile	Arg	Gln	Ala	His 295	Cys	Asn	Leu	Ser	Arg 300
	Thr	Gln	Trp	Asn	Asn 305	Thr	Leu	Gly	Lys	Ile 310	Val	Ile	Lys	Leu	Arg 315
55	Glu	Gln	Phe	Arg	Lys 320	Gln	Phe	Gly	Glu	Lys 325	Thr	Ile	Val	Phe	Asn 330
60	Arg	Ser	Ser	Gly	Gly 335	Asp	Pro	Glu	Ile	Ala 340	Met	His	Ser	Phe	Asn 345
60	Cys	Gly	Gly	Glu	Phe 350	Phe	Tyr	Cys	Asn	Thr 355	Thr	Ala	Leu	Phe	Asn 360
65	Ser	Thr	Trp	Asn	Val 365	Thr	Lys	Gly	Leu	Asn 370	Asn	Thr	Glu	Gly	Asn 375

	Se	er T	hr G	ly As	38	u As O	n Il	e Il	e Le	u Pr	o Cy:	8 Ar	g Ile	: Lys	Gln 390
5	11	e I	le As	эл Ме	t Tr ₁	p Gl: 5	n Gl	u Va	1 G1	y Ly:	s Ala	a Me	t Tyr	Ala	Pro 405
	Pr	0 I	le S∈	er Gl	y Gli 410	n Ile D	e Ar	g Cy	s Se	r Se:	. Ası	110	∋ Thr	Gly	Leu 420
10	Le	u Le	eu Th	r Ar	g Ası 429	Gly	y Gl	y Se	r Ly	s Asr 430	Glu	Se:	: Ile	Thr	Thr 435
15	Gl	u Va	al Ph	e Ar	g Pro 440	Gly	/ Gl;	y Gl	y As _l	p Met 445	Arg	Asp	neA c	Trp	Arg 450
	Se	r Gl	u Le	u Ty	r Lys 455	Туг	Lys	s Val	l Val	1 Lys 460	Ile	Glu	Pro	Leu	Gly 465
20	Va	l Al	a Pr	O Th	r Lys 470	Ala	Lys	Arç	g Arq	7 Val 475	Val	Gln	Arg	Glu	Lys 480
	Ar	g Al	a Va	1 G1;	7 Thr 485	Ile	Gly	Ala	. Met	Phe 490	Leu	Gly	Phe	Leu	Gly 495
25	Ala	a Xa	a Se	r Phe	≥ Xaa 500		Arg	Pro	Ala 504						
30	(2)	(i)	SEQU (A) : (B) :	ence Lengi Cype:	FOR CHAR CH: 1 Nuc IDEDNI	ACTE 501 leic	RIST base Aci	ICS: pai d							
35	()	i) :	(D) : SEQUI	TOPOL ENCE GTA	OGY: DESCI CCT Pro	Lin RIPT GTG	ear ION: TGG	SEQ	GAA	GCA	ACT	ACC Thr	ACT Thr	36	
40	СТА		1		TCA	5					10				
	Leu	Phe	Cys	Ala	Ser	Asp	Ala	Lys 20	Ala	Tyr	Asn	Thr	GAG Glu 25	75	
45	AAA Lys	CAT His	TAA 1 Asn	GTT Val	TGG Trp 30	GCC Ala	ACA Thr	CAC His	GCC Ala	TGT Cys 35	GTA Val	CCC Pro	ACA Thr	114	
50	GAT Asp	Pro 40) Asn	CCA Pro	CAA Gln	GAA Glu	GTA Val 45	GTA Val	TTG Leu	GGA Gly	AAT Asn	GTG Val 50	ACA Thr	153	
	GAA Glu	AAT Asn	TTT Phe	AAC Asn 55	ATG Met	TGG Trp	AAA Lys	TAA naA	AAC Asn 60	ATG Met	GTA Val	GAA Glu	CAA : Gln	192	
55	ATG Met 65	CAT His	GAA Glu	GAT Asp	ATA Ile	ATC Ile 70	AGT Ser	TTA Leu	TGG Trp	GAT (CAA i Gln : 75	AGT Ser	CTA 2 Leu	?31	
60	AAG Lys	CCA Pro	TGT Cys 80	GTA Val	AAA Lys	TTA Leu	ACC Thr	CCA Pro 85	CTC Leu	TGT (Cys \	CTT /	ACT '	TTA 2 Leu 90	?70	

	AAT Asn	TGC Cys	ACT Thr	GAT Asp	GAT Asp 95	TTA Leu	GGG Gly	AAT Asn	GCT Ala	ACT Thr 100	AAT Asn	ACC Thr	TAA neA	309
5	AGC Ser	AGT Ser 105	GCC Ala	ACT Thr	ACC Thr	AAT Asn	AGT Ser 110	AGT Ser	AGT Ser	TGG Trp	GAA Glu	GAA Glu 115	ATG Met	348
10	AAG Lys	GGG Gly	GAA Glu	ATG Met 120	AAA Lys	AGG Arg	TGC Cys	TCT Ser	TTC Phe 125	AAT Asn	ATC Ile	ACC Thr	ACA Thr	387
15	AGC Ser 130	ATA Ile	AGA Arg	GAT Asp	AAG Lys	ATT Ile 135	AAG Lys	AAA Lys	GAA Glu	CAT His	GCA Ala 140	CTT Leu	TTC Phe	426
20	TAT Tyr	AGA Arg	CTT Leu 145	GAT Asp	GTA Val	GTA Val	CCA Pro	ATA Ile 150	GAT Asp	AAT Asn	GAT Asp	AAT Asn	ACC Thr 155	465
20	ACA Thr	TAT Tyr	AGG Arg	TTG Leu	ATA Ile 160	AAT Asn	TGT Cys	AAT Asn	ACC Thr	TCA Ser 165	GTC Val	ATT Ile	ACA Thr	504
25	CAG Gln	GCC Ala 170	TGT Cys	CCA Pro	AAG Lys	GTA Val	TCA Ser 175	TTT Phe	GAG Glu	CCA Pro	ATT Ile	CCC Pro 180	ATA Ile	543
30	CAT His	TTT Phe	TGT Cys	GCC Ala 185	CCG Pro	GCT Ala	GGT Gly	TTT Phe	GCG Ala 190	ATT Ile	CTA Leu	AAG Lys	TGT Cys	582
35	AAT Asn 195	AAT Asn	AAG Lys	ACG Thr	TTC Phe	GAG Glu 200	GGA Gly	AAA Lys	GGA Gly	CCA Pro	TGT Cys 205	AAA Lys	AAT Asn	621
4.0	GTC Val	AGT Ser	ACA Thr 210	GTA Val	CAA Gln	TGC Cys	ACA Thr	CAT His 215	GGA Gly	ATT Ile	AGG Arg	CCA Pro	GTA Val 220	660
40	GTG Val	TCA Ser	ACT Thr	CAA Gln	CTG Leu 225	CTG Leu	TTA Leu	AAT Asn	GGC Gly	AGT Ser 230	CTA Leu	GCA Ala	GAA Glu	699
45	GAA Glu	GAG Glu 235	GTA Val	ATA Ile	ATT Ile	AGA Arg	TCT Ser 240	GGC Gly	TAA neA	ATC Ile	ACA Thr	GAC Asp 245	AAT Asn	738
50					ATA Ile									777
55	ATT Ile 260	AAT Asn	TGT Cys	ACA Thr	AGA Arg	TCC Ser 265	AAC Asn	AAC Asn	AAT Asn	ACA Thr	AGA Arg 270	AAA Lys	AGT Ser	816
	ATA Ile	CAT His	ATA Ile 275	GGA Gly	CCA Pro	GGG Gly	AGT Ser	GCA Ala 280	TTT Phe	TTT Phe	GCA Ala	ACA Thr	GGA Gly 285	855
60	GAA Glu	ATA Ile	ATA Ile	GGA Gly	GAT Asp 290	ATA Ile	AGA Arg	CAA Gln	GCA Ala	CAC His 295	TGT Cys	AAC Asn	CTT Leu	894

	AG' Se:	r Ar 30	g Thi	A CAI	A TGG	AA! Asi	n Ası 30	n Th	T TT. r Le	A GG	A AAG y Lys	TATA Ile 310	• Val	933
5	AT	A AA ≥ Ly:	A TTA	A AGA Arg 319	g Glu	CA!	A TT	P AG ⊇ Ar	A AA g Ly: 320	s Gl	A TTT	r GGA e Gly	GAA Glu	972
10	AAJ Lys 325	Th:	A ATA	GTC Val	TTI Phe	AA1 Asr 330	n Arg	TC:	C TC/ r Se	A GGA	GGC Gly 335	Asp	CCG Pro	101:
15	GA <i>F</i> Glu	A ATT	T GCA ∋ Ala 340	Met	CAC His	AGT Ser	TT1	AA: Asi 349	n Cys	r GGA Gly	GGG Gly	GAA Glu	TTT Phe 350	1050
20	TTC Phe	TAC Tyr	Cys	AAC Asn	ACA Thr 355	Thr	GCA Ala	CTO Lev	3 TT1 1 Phe	TAA : Asn 360	Ser	ACC Thr	TGG Trp	1089
20	AAT Asn	GT1 Val 365	. Thr	AAA Lys	GGG Gly	TTG Leu	AAT Asn 370	Asr	C ACT	GAA Glu	GGA Gly	AAT Asn 375	AGC Ser	1128
25	ACA Thr	GGG	GAT Asp	GAA Glu 380	Asn	ATC Ile	ATA Ile	CTC Leu	CCA Pro 385	Cys	AGA Arg	ATA Ile	AAA Lys	1167
30	CAA Gln 390	Ile	ATA	AAC Asn	ATG Met	TGC Trp 395	CAG Gln	GAA Glu	GTA Val	GGA Gly	AAA Lys 400	GCA Ala	ATG Met	1206
35	TAT Tyr	GCC Ala	CCT Pro 405	CCC Pro	ATC Ile	AGT Ser	GGA Gly	CAA Gln 410	Ile	AGA Arg	TGT Cys	TCA Ser	TCA Ser 415	1245
40	TAA neA	ATT	ACA Thr	GGG Gly	CTG Leu 420	CTA Leu	CTA Leu	ACA Thr	AGA Arg	GAT Asp 425	GGT Gly	GGT Gly	AGT Ser	1284
40	AAG Lys	AAC Asn 430	GAG Glu	AGC Ser	ATC Ile	ACC Thr	ACC Thr 435	GAG Glu	GTC Val	TTC Phe	AGA Arg	CCT Pro 440	GGA Gly	1323
45	GGA Gly	GGA Gly	GAT Asp	ATG Met 445	AGG Arg	GAC Asp	AAT Asn	TCG Trp	AGA Arg 450	AGT Ser	GAA Glu	TTA Leu	TAT Tyr	1362
50	AAA Lys 455	TAT Tyr	AAA Lys	GTA Val	GTA Val	AAA Lys 460	ATT Ile	GAA Glu	CCA Pro	TTA Leu	GGA Gly 465	GTA Val	GCG Ala	1401
5 5	CCC Pro	ACC Thr	AAG Lys 470	GCA Ala	AAG Lys	AGA Arg	AGA Arg	GTG Val 475	GTG Val	CAG Gln	AGA Arg	Glu	AAA Lys 480	1440
60	AGA Arg	GCA Ala	GTG Val	GGA Gly	ACA Thr 485	ATA Ile	GGA Gly	GCT Ala	ATG Met	TTC Phe 490	CTT Leu	GGG (TTC :	1479
60			GCA Ala					A 15	501					

```
(2) INFORMATION FOR SEQ ID NO:12:
         (i) SEQUENCE CHARACTERISTICS:
             (A) LENGTH: 500 amino acids
             (B) TYPE: Amino Acid
 5
             (D) TOPOLOGY: Linear
        (xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:
      Leu Glu Val Pro Val Trp Lys Glu Ala Thr Thr Thr Leu Phe Cys
10
      Ala Ser Asp Ala Lys Ala Tyr Asn Thr Glu Lys His Asn Val Trp
      Ala Thr His Ala Cys Val Pro Thr Asp Pro Asn Pro Gln Glu Val
15
      Val Leu Gly Asn Val Thr Glu Asn Phe Asn Met Trp Lys Asn Asn
      Met Val Glu Gln Met His Glu Asp Ile Ile Ser Leu Trp Asp Gln
20
      Ser Leu Lys Pro Cys Val Lys Leu Thr Pro Leu Cys Val Thr Leu
25
       Asn Cys Thr Asp Asp Leu Gly Asn Ala Thr Asn Thr Asn Ser Ser
      Ala Thr Thr Asn Ser Ser Ser Trp Glu Glu Met Lys Gly Glu Met
30
                                           115
      Lys Arg Cys Ser Phe Asn Ile Thr Thr Ser Ile Arg Asp Lys Ile
                                           130
      Lys Lys Glu His Ala Leu Phe Tyr Arg Leu Asp Val Val Pro Ile
35
      Asp Asn Asp Asn Thr Thr Tyr Arg Leu Ile Asn Cys Asn Thr Ser
                                           160
40
       Val-Ile Thr Gln Ala Cys Pro Lys Val Ser Phe Glu Pro Ile Pro
                                       . 175
       Ile His Phe Cys Ala Pro Ala Gly Phe Ala Ile Leu Lys Cys Asn
                                           190
45
      Asn Lys Thr Phe Glu Gly Lys Gly Pro Cys Lys Asn Val Ser Thr
       Val Gln Cys Thr His Gly Ile Arg Pro Val Val Ser Thr Gln Leu
50
                                           220
       Leu Leu Asn Gly Ser Leu Ala Glu Glu Glu Val Ile Ile Arg Ser
55
       Gly Asn Ile Thr Asp Asn Thr Lys Thr Ile Ile Val Gln Leu Asn
       Glu Ser Val Val Ile Asn Cys Thr Arg Ser Asn Asn Asn Thr Arg
60
       Lys Ser Ile His Ile Gly Pro Gly Ser Ala Phe Phe Ala Thr Gly
```

	Glu Ile Ile Gl	y Asp Ile 290	e Arg Gln	Ala Hi 29	s Cys As 5	n Leu Ser Arg 300
5	Thr Gln Trp As	n Asn Thr 305	Leu Gly	Lys Ile		e Lys Leu Arg 315
	Glu Gln Phe Ar	g Lys Gln 320	Phe Gly	Glu Lys 325	Thr Il	e Val Phe Asn 330
10	Arg Ser Ser Gl	y Gly Asp 335	Pro Glu	Ile Ala	a Met Hi)	s Ser Phe Asn 345
15	Cys Gly Gly Gl	Phe Phe 350	Tyr Cys	Asn Thr 355		a Leu Phe Asn 360
	Ser Thr Trp As	val Thr 365	Lys Gly	Leu Asn 370		Glu Gly Asn 375
20	Ser Thr Gly Asp	Glu Asn 380	4le Ile	Leu Pro 385		Ile Lys Gln 390
	Ile Ile Asn Met	Trp Gln 395	Glu Val	Gly Lys 400		Tyr Ala Pro 405
25	Pro Ile Ser Gly	Gln Ile 410	Arg Cys	Ser Ser 415	Asn Ile	Thr Gly Leu 420
30	Leu Leu Thr Arc	Asp Gly 425	Gly Ser	Lys Asn 430	Glu Ser	Ile Thr Thr 435
	Glu Val Phe Arg	Pro Gly 440	Gly Gly	Asp Met 445	Arg Asp	Asn Trp Arg 450
35	Ser Glu Leu Tyr	Lys Tyr 455	Lys Val	Val Lys 460	Ile Glu	Pro Leu Gly 465
	Val Ala Pro Thr	Lys Ala 470	Lys Arg	Arg Val 475	Val Gln	Arg Glu Lys 480
40	Arg Ala Val Gly	Thr Ile 485	Gly Ala	Met Phe 490	Leu Gly	Phe Leu Gly 495
45	Ala Xaa Ser Phe	Xaa 500				
	(2) INFORMATION (i) SEQUENCE (A) LENGT	CHARACTER	ISTICS:	a		
50	(B) TYPE: (C) STRAN (D) TOPOLO	Nucleic DEDNESS:	Acid Single	3		
	(xi) SEQUENCE	DESCRIPTION	ON: SEQ	ID NO:13	:	
55	GG GAA TTC GGA Glu Phe Gly 1	TCC GGG (Ser Gly)	GTA CCT (Val Pro V	STG TGG Val Trp	AAG GAA Lys Glu 10	GCA 38 Ala
60	ACC ACC ACT CTA Thr Thr Thr Leu 15	TTC TGT (GCA TCA (Ala Ser A 20	GAT GCT Asp Ala	AGA GCA Arg Ala	TAT 77 Tyr 25
65	GAC ACA GAG GTA Asp Thr Glu Val	CAT AAT (His Asn \ 30	GTT TGG (Val Trp #	SCC ACA Ala Thr 35	CAT GCC His Ala	TGT 116 Cys

	GTA Val	CCC Pro 40	ACA Thr	GAC Asp	CCT Pro	AGT Ser	CCA Pro 45	CAA Gln	GAA Glu	GTA Val	GTT Val	TTG Leu 50	GAA Glu	155
5	AAT Asn	GTG Val	ACA Thr	GAA Glu 55	AAT Asn	TTT Phe	AAC Asn	ATG Met	TGG Trp 60	AAA Lys	AAT Asn	AAC Asn	ATG Met	194
10	Val 65	Glu	Gln	Met	His	Glu 70	Asp	Ile	Ile	Ser	Leu 75	Trp		
15	Gln	Ser	Leu 80	Lys	Pro	Cys	Val	Lys 85	Leu	Thr	Pro	Leu	90	
	GTT Val	ACT Thr	TTA Leu	AAT Asn	TGC Cys 95	AGT Ser	GAT Asp	TAT Tyr	AGG Arg	AAT Asn 100	GCT Ala	ACT Thr	GAT Asp	311
20	TAT Tyr	AAG Lys 105	AAT Asn	GCT Ala	ACT Thr	GAT Asp	ACC Thr 110	ACT Thr	AGT Ser	AGT Ser	AAC Asn	GAG Glu 115	GGA Gly	350
25	AAG Lys	ATG Met	GAG Glu	AGA Arg 120	GGA Gly	GAA Glu	ATA Ile	AAA Lys	AAC Asn 125	TGC Cys	TCT Ser	TTC Phe	AAT Asn	389
30	ATT Ile 130	ACC Thr	ACA Thr	AGC Ser	ATA Ile	AAA Lys 135	AAT Asn	AAG Lys	ATG Met	CAG Gln	AAA Lys 140	GAA Glu	TAT Tyr	428
35	GCA Ala	CTT Leu	TTC Phe 145	TAT Tyr	AAA Lys	CTT Leu	GAT Asp	ATA Ile 150	GTA Val	CCA Pro	ATA Ile	GAT Asp	AAT Asn 155	467
	ACA Thr	AGC Ser	TAT Tyr	ACA Thr	TTG Leu 160	ATA Ile	AGT Ser	TGT Cys	AAC Asn	ACC Thr 165	TCA Ser	GTC Val	ATT Ile	506
40	ACA Thr	CAG Gln 170	GCC Ala	TGT Cys	CCA Pro	AAG Lys	GTA Val 175	TCC Ser	TTT Phe	GAA Glu	CCA Pro	ACT Thr 180	CCC Pro	545
45	ATA Ile	CAT His	TAT Tyr	TGT Cys 185	GCT Ala	CCG Pro	GCT Ala	GGT Gly	TTT Phe 190	GCG Ala	ATT Ile	CTA Leu	AAG Lys	584
50	TGT Cys 195	AAT Asn	GAT Asp	AAG Lys	AAG Lys	TTC Phe 200	AGT Ser	GGA Gly	AAA Lys	GGA Gly	GAA Glu 205	TGT Cys	AAA Lys	623
55	AAT Asn	GTC Val	AGC Ser 210	ACA Thr	GTA Val	CAA Gln	TGT Cys	ACA Thr 215	CAT His	GGA Gly	ATT Ile	AGG Arg	CCA Pro 220	662
	GTA Val	GTA Val	TCA Ser	ACT Thr	CAA Gln 225	CTG Leu	CTG Leu	TTA Leu	AAT Asn	GGC Gly 230	AGT Ser	CTA Leu	GCA Ala	701
60	GAA Glu	GAA Glu 235	GAG Glu	GTG Val	GTA Val	ATT Ile	AGA Arg 240	TCT Ser	GAC Asp	AAT Asn	TTC Phe	ATA Ile 245	GAC Asp	740

	AAT Asn	ACT Thr	'AAA 'Lys	ACC Thr 250	Ile	ATA Ile	GTA Val	A CAG	G CTC n Lec 259	ı Lys	A GAA	TC' Se	r GT/ r Val	A 779
5	GAA Glu 260	Ile	AAT Asn	TGT Cys	ATA Ile	AGA Arg 265	Pro	AAC Ası	C AAT n Asr	C AAT n Asr	ACA Thr 270	Arg	A AA/	818
10	GGT Gly	ATA Ile	CAT His 275	Ile	GGA Gly	CCA Pro	GGG	AGA Arc 280	g Ala	TGC Trp	TAT Tyr	GCA Ala	A ACA Thr 285	
15	GGA Gly	GAA Glu	ATA Ile	GTA Val	GGA Gly 290	Asp	ATA Ile	AG#	A AAG J Lys	GCA Ala 295	Tyr	TGT	AAC Asn	896
20	ATT Ile	AGT Ser 300	AGA Arg	ACA Thr	AAA Lys	TGG Trp	AAT Asn 305	AAC	ACT Thr	TTA Leu	λTA Ile	CAG Gln 310	Ile	935
	GCT Ala	AAC Asn	AAA Lys	TTA Leu 315	AAA Lys	GAA Glu	AAA Lys	TAT Tyr	AAT Asn 320	Thr	ACA Thr	ATA Ile	AGC Ser	974
25	TTT Phe 325	AAT Asn	CGA Arg	TCC Ser	TCA Ser	GGA Gly 330	GGG Gly	GAC Asp	CCA Pro	GAA Glu	ATT Ile 335	GTA Val	ACG Thr	1013
30	CAT His	AGT Ser	TTT Phe 340	AAT Asn	TGT Cys	GGA Gly	GGG Gly	GAG Glu 345	TTT Phe	TTC Phe	TAC Tyr	TGT Cys	GAT Asp 350	1052
35	TCA Ser	ACA Thr	CAA Gln	CTG Leu	TTT Phe 355	AAT Asn	AGT Ser	ACT Thr	TGG Trp	AAT Asn 360	TTA Leu	AAT Asn	GGT Gly	1091
40	ACT Thr	TGG Trp 365	AAT Asn	TTT Phe	ACT Thr	GCA Ala	GGG Gly 370	TCA Ser	AAT Asn	GAA Glu	ACT Thr	GAA Glu 375	GGC	1130
	AAT Asn	ATC Ile	ACA Thr	CTC Leu 380	CCA Pro	TGC Cys	AGA Arg	ATA Ile	AAA Lys 385	CAA Gln	ATT Ile	ATA Ile	AAC Asn	1169
45	AGG Arg 390	TGG Trp	CAG Gln	GAA (Glu	Val	GGG Gly 395	AAA Lys	GCA Ala	ATG Met	TAT Tyr	GCC Ala 400	CCT Pro	CCC Pro	1208
50	ATC Ile	Ser	GGA Gly 405	CAA /	ATA . Ile	AAA Lys	Cys	TCA Ser 410	TCA Ser	AAC Asn	ATT Ile	ACA Thr	GGG Gly 415	1247
55	ATG . Met	ATA (TTA : Leu '	The A	AGG (Arg (GAT (GGT Gly	GGT Gly	Asn	GAG Glu 425	AAC Asn	AAT Asn	TAA neA	1286
60	GAG A	AGC 1 Ser 3 430	AGT :	ACT I	ACT (Glu 1	ACC Thr 435	TTC Phe	AGA Arg	CCG (Gly (GGA Gly 440	GGA Gly	1325
	GAT A	ATG /	Arg /	AAC A Asn A 445	AT 1	rgg / Frp /	AGA A	Ser	GAA Glu i 450	TTA ' Leu '	TAT A	AAA Lys	TAT :	1364

		GTA Val)3
_	455			-		460					465				_
5		GCA Ala													12
10		GGA Gly													1
15		TAA Xaa 495										151	4		
20	((i	EQUEI A) LI B) T'	NCE (ENGT! (PE: DPOL(CHARA 1: 50 Amir OGY:	ACTEI 04 ar no Ac Line	RIST: mino cid ear	ICS: acid	ds	NO: 1	4 :				
25	Glu 1	Phe	Gly	Ser	Gly 5	Val	Pro	Val	Trp	Lys 10	Glu	Ala	Thr	Thr	Thr 15
	Leu	Phe	Cys	Ala	Ser 20	Asp	Ala	Arg	Ala	Tyr 25	Asp	Thr	Glu	Val	His 30
30	Asn	Val	Trp	Ala	Thr 35	His	Ala	Cys	Val	Pro 40	Thr	Asp	Pro	Ser	Pro 45
35	Gln	Glu	Val	Val	Leu 50	Glu	Asn	Val	Thr	Glu 55	Asn	Phe	Asn	Met	Trp 60
3 3	Lys	Asn	Asn	Met	Val 65	Glu	Gln	Met	His	Glu 70	Asp	Ile	lle	Ser	Leu 75
40	Trp	Asp	Gln	Ser	Leu 80	Lys	Pro	Cys	Val	Lys 85	Leu	Thr	Pro	Leu	Cys 90
	Val	Thr	Leu	Asn	Cys 95	Ser	Aśp	Tyr	Arg	Asn 100	Ala	Thr	Asp	Tyr	Lys 105
45	Asn	Ala	Thr	Asp	Thr 110	Thr	Ser	Ser	Asn	Glu 115	Gly	Lys	Met	Glu	Arg 120
50	Gly	Glu	Ile	Lys	Asn 125	Cys	Ser	Phe	Asn	Ile 130	Thr	Thr	Ser	Ile	Lys 135
	Asn	Lys	Met	Gln	Lys 140	Glu	Tyr	Ala	Leu	Phe 145	Tyr	Lys	Leu	Asp	Ile 150
55	Val	Pro	Ile	Asp	Asn 155	Thr	Ser	Tyr	Thr	Leu 160	Ile	Ser	Cys	neA	Thr 165
	Ser	Val	Ile	Thr	Gln 170	Ala	Cys	Pro	Lys	Val 175	Ser	Phe	Glu	Pro	Thr 180
60	Pro	Ile	His	Tyr	Cys 185	Ala	Pro	Ala	Gly	Phe 190	Ala	Ile	Leu	Lys	Cys 195
65	Asn	Asp	Lys	Lys	Phe 200	Ser	Gly	Lys	Gly	Glu 205	Cys	Lys	Asn	Val	Ser 210

	Thr	. Val	Gln	Cys	Thr 215		Gly	Ile	e Arq	220		. Val	l Se	r Thi	Gln 225
5	Leu	Leu	Leu	Asn	Gly 230		Leu	Ala	Glu	Glu 235		Val	l Va	l Ile	240
	Ser	Asp	Asn	Phe	11e 245	Asp	Asn	Thr	Lys	Thr 250		Ile	e Val	l Glr	Leu 255
10	Lys	Glu	Ser	Val	Glu 260	Ile	Asn	Cys	Ile	Arg 265		Asn	Asr) Asn	Thr 270
15					275					280					Thr 285
	Gly	Glu	Ile	Val	Gly 290	Asp	Ile	Arg	Lys	Ala 295	Tyr	Cys	Asn	lle	Ser 300
20					305					310				Lys	315
					320					325				Ser	330
25					335					340			_	Gly	345
30					350					355				Trp	360
					365					370				Thr	375
35					380					385				Asn	39Ō
					395					400				Ser	405
40					410					415				Thr	420
45					425					430				Glu	435
					440					445			-	Ser	450
50	Leu	Tyr	Lys	Tyr	Lys 455	Val	Val	Lys	Ile	Glu 460	Pro	Leu	Gly	Val	Ala 465
					470					475			-	Arg	480
55	Val	Gly	Ala	Leu	Gly 485	Ala i	Met	Phe		Gly 490	Phe	Leu	Gly	Ala	Xaa 495
60	Ser (2) I				500	Ser i		-	Ser 504						
) SE	QUEN	CE CI	HARA	CTER: 08 ba	ISTI ase	CS:	s						
65						ss: s		le							

-120-

(D) TOPOLOGY: Linear (xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

5	G	GTA Val 1	CCT Pro	GTG Val	TGG Trp	AAG Lys 5	GAA Glu	GCA Ala	ACC Thr	ACC Thr	ACT Thr 10	CTA Leu	TTC Phe	37
10	TGT Cys	GCA Ala	TCA Ser 15	GAT Asp	GCT Ala	AGA Arg	GCA Ala	TAT Tyr 20	GAC Asp	ACA Thr	GAG Glu	GTA Val	CAT His 25	76
	AAT Asn	GTT Val	TGG Trp	GCC Ala	ACA Thr 30	CAT His	GCC Ala	TGT Cys	GTA Val	CCC Pro 35	ACA Thr	GAC Asp	CCT Pro	115
15	AGT Ser	CCA Pro 40	CAA Gln	GAA Glu	GTA Val	TTT Phe	TTG Leu 45	GGA Gly	AAT Asn	GTG Val	ACA Thr	GAA Glu 50	AAT Asn	154
20	TTT Phe	AAT Asn	ATG Met	TGG Trp 55	FÀ2 YYY	AAT Asn	AAC Asn	ATG Met	GTA Val 60	GAA Glu	CAA Gln	ATG Met	TAT Tyr	193
25	GAG Glu 65	GAT Asp	ATA Ile	ATT Ile	AGT Ser	TTA Leu 70	TGG Trp	GAT Asp	CAA Gln	AGC Ser	TTA Leu 75	AAG Lys	CCA Pro	232
30	TGT Cys	GTA Val	AAA Lys 80	TTA Leu	ACC Thr	CCA Pro	CTC Leu	TGT Cys 85	GTT Val	ACT Thr	TTA Leu	AAT Asn	TGC Cys 90	271
	AGT Ser	GAT Asp	TAT Tyr	AGG Arg	AAT Asn 95	GCT Ala	ACT Thr	GAT Asp	TAT Tyr	AAG Lys 100	AAT Asn	GCT Ala	ACT Thr	310
35	GAT Asp	ACC Thr 105	ACT Thr	AGT Ser	AGT Ser	AAC Asn	GAG Glu 110	GGA Gly	AAG Lys	ATG Met	GAG Glu	AGA Arg 115	GGA Gly	349
40	GAA Glu	ATA Ile	AAA Lys	AAC Asn 120	TGC Cys	TCT Ser	TTC Phe	AAT Asn	ATC Ile 125	ACC Thr	ACA Thr	AGC Ser	ATA Ile	388
45	AAA Lys 130	AAT Asn	AAG Lys	ATG Met	CAG Gln	AAA Lys 135	GAA Glu	TAT Tyr	GCA Ala	CTT Leu	TTC Phe 140	TAT Tyr	AAA Lys	427
50	CTT Leu	AAT Asn	ATA Ile 145	GTA Val	CCA Pro	ATA Ile	GAT Asp	AAT Asn 150	ACA Thr	AGC Ser	TAT Tyr	ACA Thr	TTG Leu 155	466
	ATA Ile	AGT Ser	TGT Cys	AAC Asn	ACC Thr 160	TCA Ser	GTC Val	ATT Ile	ACA Thr	CAG Gln 165	GCC Ala	TGT Cys	CCA Pro	505
55	AAG Lys	GTA Val 170	TCC Ser	TTT Phe	GAA Glu	CCA Pro	ATT Ile 175	CCC Pro	ATA Ile	CAT His	TAT Tyr	TGT Cys 180	GCT Ala	544
60	CCG Pro	GCT Ala	GGT Gly	TTT Phe 185	GCG Ala	ATT Ile	CTA Leu	AAG Lys	TGT Cys 190	AAT Asn	GAT Asp	AAG Lys	AAG Lys	583

	TT(Phe 199	Sei	r GGA r Gly	A AAA / Lys	GGA Gly	GAA Glu 200	ı Cys	T AAA E Lys	A AAT B Ast	CTC Val	Ser 205	Thr	GTA Val	622
5	CAA Glr	TG1	T ACA Thr 210	. His	GGA Gly	ATI	AGC Arg	Pro 215	Val	GTA Val	TCA Ser	ACT Thr	CAA Gln 220	
10	CTG Leu	CTC Lev	TTA Leu	AAT Asn	GGC Gly 225	Ser	CTA Leu	GCA Ala	A GAA I Glu	GAA Glu 230	Glu	GTG Val	GTA Val	700
15	ATT Ile	AGA Arg 235	, Ser	GAC Asp	AAT Asn	TTC Phe	Thr 240	Asp	AAT Asn	ACT Thr	AAA Lys	ACC Thr 245	ATA Ile	739
20	ATA Ile	GTA Val	CAG Gln	CTG Leu 250	Lys	GAA Glu	TCT Ser	GTA Val	GAA Glu 255	Ile	AAT Asn	TGT Cys	ATA Ile	778
20	AGA Arg 260	Pro	AAC Asn	AAT Asn	AAT Asn	ACA Thr 265	AGA Arg	AAA Lys	GGT	ATA Ile	CAT His 270	ATA Ile	GGA Gly	817
25	CCA Pro	GGG Gly	AGA Arg 275	GCA Ala	TGG Trp	TAT Tyr	GCA Ala	ACA Thr 280	Gly	GAA Glu	ATA Ile	GTA Val	GGA Gly 285	856
30	GAT Asp	ATA Ile	AGA Arg	CAG Gln	GCA Ala 290	TAT Tyr	TGT Cys	AAC Aan	ATT	AGT Ser 295	AGA Arg	ACA Thr	AAA Lys	895
35	TGG Trp	AAT Asn 300	AAC Asn	ACT Thr	TTA Leu	ATA Ile	CAG Gln 305	ATA Ile	GCT Ala	AAC Asn	AAA Lys	TTA Leu 310	AAA Lys	934
40	GAA Glu	AAA Lys	TAT Tyr	AAT Asn 315	ACA Thr	ACA Thr	ATA Ile	AGC Ser	TTT Phe 320	AAT Asn	CGA Arg	TCC Ser	TCA Ser	973
40	GGA Gly 325	GGG Gly	GAC Asp	CCA Pro	GAA Glu	ATT Ile 330	GTA Val	ACC Thr	CAT His	AGT Ser	TTT Phe 335	AAT Asn	TGT Cys	1012
45	GGA Gly	GGG Gly	GAA Glu 340	TTT Phe	TTC Phe	TAC Tyr	TGT Cys	AAT Asn 345	TCA Ser	ACA Thr	CAA Gln	CTG Leu	TTT Phe 350	1051
50	AAT Asn	AGT Ser	ACT Thr	TGG Trp	AAT Asn 355	TTA Leu	AAT Asn	GGT Gly	ACT Thr	TGG Trp 360	AAT Asn	TTT Phe	ACT Thr	1090
55	GCA Ala	GGG Gly 365	TCA Ser	AAT Asn	GAA Glu	ACT Thr	GAA Glu 370	GGC Gly	AAT Asn	ATC Ile	Thr	CTC Leu 375	CCA Pro	1129
60	TGC Cys	AGA Arg	ATA Ile	AAA Lys 380	CAA Gln	ATT Ile	ATA Ile	AAC Asn	AGG Arg 385	TGG Trp	CAG (Gln (GAA Glu	GTA Val	1168
60	GGA Gly 390	AAA Lys	GCA Ala	ATG Met	Tyr	GCC Ala 395	CCT Pro	CCC Pro	ATC Ile	Ser	GGA (Gly (CAA . Gln	ATA : Ile	1207

	AGA Arg	TGC	TCA Ser 405	TCA Ser	AAC Asn	ATT Ile	ACA Thr	GGG Gly 410	ATG Met	ATA Ile	TTA Leu	ACA Thr	AGG Arg 415	124	6
5	GAT Asp	GGT Gly	GGT Gly	AAC Asn	GAG Glu 420	AAC Asn	AAT Asn	AAT Asn	GAG Glu	AGC Ser 425	AGT Ser	ACT Thr	ACT Thr	128	5
10	GAG Glu	ACC Thr 430	TTC Phe	AGA Arg	CCG Pro	GGA Gly	GGA Gly 435	GGA Gly	GAT Asp	ATG Met	AGG Arg	AAC Asn 440	AAT Asn	132	4
15	TGG Trp	AGA Arg	AGT Ser	GAA Glu 445	TTA Leu	TAT Tyr	AAA Lys	TAT Tyr	AAA Lys 450	GTA Val	GTA Val	AAA Lys	ATT Ile	136.	3
20	GAG Glu 455	CCA Pro	TTA Leu	GGA Gly	GTA Val	GCA Ala 460	CCC Pro	ACC Thr	GAC Asp	TCT Ser	AGA Arg 465	GGA Gly	TCC Ser	140	2
20	TCT Ser	AGA Arg 469	1408	3											
25	(2) I (i	.) SE (<i>P</i> (E	EQUEN () LE 3) TY	NCE C ENGTH (PE:	CHARA i: 46 Amir	CTER 69 am 10 Ac	RISTI nino cid	cs:							
30	(xi	() .) SE	OUE)	ICE I	OGY: DESCF	RIPTI	ON:	SEQ	ID N	10:16	;				
	Val 1	Pro	Val	Trp	Lys 5	Glu	Ala	Thr	Thr	Thr 10	Leu	Phe	Cys	Ala	Ser 15
35	Asp	Ala	Arg	Ala	Tyr 20	Asp	Thr	Glu	Val	His 25	Asn	Val	Trp	Ala	Thr 30
40	His	Ala	Cys	Val	Pro 35	Thr	Asp	Pro	Ser	Pro 40	Gln	Glu	Val	Phe	Leu 45
40	Gly	Asn	Val	Thr	Glu 50	Asn	Phe	Asn	Met	Trp 55	Lys	Asn	Asn	Met	Val 60
45	Glu	Gln	Met	Tyr	Glu 65	Asp	Ile	Ile	Ser	Leu 70	Trp	Asp	Gln	Ser	Leu 75
	Lys	Pro	Cys	Val	Lys 80	Leu	Thr	Pro	Leu	Cys 85	Val	Thr	Leu	Asn	Cys 90
50	Ser	Asp	Tyr	Arg	Asn 95	Ala	Thr	Asp	Tyr	Lys 100	Asn	Ala	Thr	Asp	Thr 105
	Thr	Ser	Ser	Asn	Glu 110	Gly	Lys	Met	Glu	Arg 115	Gly	Glu	Ile	Lys	Asn 120
55	Cys	Ser	Phe	Asn	11e 125	Thr	Thr	Ser	Ile	Lys 130	Asn	Lys	Met	Gln	Lys 135
60	Glu	Tyr	Ala	Leu	Phe 140	Tyr	Lys	Leu	Asn	Ile 145	Val	Pro	Ile	Asp	Asn 150
	Thr	Ser	Tyr	Thr	Leu 155	Ile	Ser	Cys	Asn	Thr 160	Ser	Val	Ile	Thr	Gln 165

	Ala	а Су	s Pr	o Ly	s Val	l Sei	: Phe	e Gl	a Pro	7 Ile 175	Pro	o Ile	e Hi	s Ty	r Cys 180
5	Ala	Pr	o Al	a Gl	y Phe 185	Ala	Ile	e Le	ı Lys	3 Cys	Asr	Asp	Ly	s Ly	s Phe 195
	Ser	Gl	y Ly:	s Gl	y Glu 200	Cys	Lys	A A s r	val	Ser 205	Thr	Val	Gl	n Cy	s Thr 210
10	His	Gly	y Ile	e Aro	Pro 215	Val	Val	Ser	Thr	Gln 220	Leu	Leu	Lei	u Ası	n Gly 225
15	Ser	Le	ı Ala	a Glu	Glu 230	Glu	Val	Val	Ile	Arg 235	Ser	Asp	Ası	n Phe	Thr 240
	Asp	Asr	Thr	Lys	Thr 245	Ile	Ile	Val	Gln	Leu 250	Lys	Glu	Ser	· Val	Glu 255
20	Ile	Asn	Cys	: Ile	Arg 260	Pro	Asn	Asn	Asn	Thr 265	Arg	Lys	Gly	Ile	His 270
	Ile	Gly	Pro	Gly	Arg 275	Ala	Trp	Tyr	Ala	Thr 280	Gly	Glu	Ile	val	Gly 285
25	Asp	Ile	Arg	Gln	Ala 290	Tyr	Cys	Asn	Ile	Ser 295	Arg	Thr	Lys	Trp	Asn 300
30					Gln 305					310				_	315
	Thr	Thr	Ile	Ser	Phe 320	Asn	Arg	Ser	Ser	Gly 325	Gly	Asp	Pro	Glu	Ile 330
35	Val	Thr	His	Ser	Phe 335	Asn	Cys	Gly	Gly	Glu 340	Phe	Phe	Tyr	Cys	Asn 345
					Phe 350					355					360
40					Gly 365					370					375
45					Lys 380					385					390
					Ala 395					400				•	405
50					Gly 410					415			-		420
	Asn i				425				•	430				_	435
55	Gly i	Asp	Met	Arg	Asn 7	Asn '	rp	Arg	Ser (31u I 445	Leu 1	Tyr 1	_ys	Tyr	Lys 450
60	Val V	Val	Lys	Ile	Glu 1 455	Pro 1	Leu (Sly '	Val A	Ala F 160	ro 1	Thr A	Asp .		Arg 465
	Gly S	Ser		Arg 469											
65	(2) IN	VFOR	MATI	ON F	OR SE	Q II	NO:	17:							

(2) INFORMATION FOR SEQ ID NO:17: (i) SEQUENCE CHARACTERISTICS:

		(1	B) T'	engti Ype: Trani	Nuc DEDNI	leic ESS:	Acie Sin	d:	rs					
5	(x.	() i) S	D) TO EQUE	OPOLO NCE I	OGY: DESCI	Line RIPT	ear ION:	SEQ	ID I	NO: 1	7:			
		GAG Glu 1	GTA Val	CCT Pro	GTG Val	TGG Trp 5	AAA Lys	GAA Glu	GCA Ala	ACC Thr	ACT Thr 10	ACT Thr	CTA Leu	36
10	TTT Phe	TGT Cys	GCA Ala 15	TCA Ser	GAT Asp	GCT Ala	AAA Lys	GCA Ala 20	TAT Tyr	GAC Asp	ACA Thr	GGG Gly	GTG Val 25	75
15	CAT His	AAT Asn	GTT Val	TGG Trp	GCC Ala 30	ACA Thr	CAT His	GCC Ala	TGT Cys	GTA Val 35	CCC Pro	ACA Thr	GAC Asp	114
20	CCC Pro	AAC Asn 40	CCA Pro	CAA Gln	GAA Glu	ATA Ile	GAA Glu 45	TTG Leu	GTA Val	AAT Asn	GTG Val	ACA Thr 50	GAA Glu	153
25	GAT Asp	TTT Phe	AAC Asn	ATG Met 55	TGG Trp	AAA Lys	AAT Asn	AAA Lys	ATG Met 60	GTA Val	GAC Asp	CAG Gln	ATG Met	192
30	CAT His 65	GAG Glu	GAT Asp	ATA Ile	ATC Ile	AGT Ser 70	TTA Leu	TGG Trp	GAT Asp	GAA Glu	AGC Ser 75	CTA Leu	AAG Lys	231
30	CCA Pro	TGT Cys	GTA Val 80	AAG Lys	TTA Leu	ACC Thr	CCA Pro	CTT Leu 85	TGT Cys	GTT Val	ACT Thr	CTA Leu	AAC Asn 90	270
35	TGC Cys	AGT Ser	GAT Asp	GTG Val	AAC Asn 95	AAT Asn	TCC Ser	ACA Thr	AAT Asn	CCT Pro 100	AAT Asn	GAT Asp	ACT Thr	309
40	AAT Asn	ACT Thr 105	AAT Asn	TCC Ser	ACT Thr	AAT Asn	ACT Thr 110	ACT Thr	TCC Ser	TCT Ser	ACT Thr	CCT Pro 115	ACG Thr	348
45	GCC Ala	ACT Thr	ACT Thr	AGT Ser 120	AGC Ser	GAG Glu	GAA Glu	AAG Lys	ATG Met 125	GAG Glu	AAG Lys	GGA Gly	GAA Glu	387
5 0	ATA Ile 130	AAA Lys	AAC Asn	TGC Cys	TCT Ser	TTC Phe 135	AAT Asn	ATC Ile	ACC Thr	ACA Thr	CAC His 140	ATG Met	AAA Lys	426
50	GAT GAT	AAG Lys	GCA Ala 145	CAG Gln	AAA Lys	GAA Glu	TAT Tyr	GCA Ala 150	CTT Leu	TTT Phe	TAT Tyr	AAA Lys	CTT Leu 155	465
55	Aab	ATA Ile	GTA Val	CCA Pro	ATA Ile 160	GAT Asp	GAT Asp	AAT Asn	AAT Asn	GCC Ala 165	AGC Ser	TAT Tyr	AGG Arg	504
60	TTG Leu	ATA Ile 170	AGT Ser	TGT Cys	AAT Asn	ACC Thr	TCA Ser 175	GAC Asp	ATT Ile	ACA Thr	CAG Gln	GCC Ala 180	TGT Cys	543
65	CCA Pro	AAG Lys	GTG Val	ACC Thr 185	TTT Phe	GAG Glu	CCA Pro	ATT Ile	CCC Pro 190	ATA Ile	CAT His	TAT Tyr	TGT Cys	582

	G(A) 19	aP	CG G	CT G	GT T ly P	ne A	CG la 00	AT:	r ct ≥ Le	A AA u Ly	G To	/s L	AA ys 05	GAT Asp	AA Ly:	G 621
5	Ly	s P	ne A 2	sn G 10	ly T	hr G	ly	Pro	21	s Se 5	r Ly	's V	al	Ser	Th: 220)
10	GT Va	A C	AA T ln C	GT A	nr H	AT G is G 25	GA ly	ATT Ile	AG Ar	G CC g Pr	A G1 O Va 23	1 V	TA al	TCA Ser	ACT Thr	699
15	CA G1	A Ci n Le 23	eu L	TG Ti	TA AA eu As	AT G	GC ly	AGT Ser 240	Lei	r GC	A GA a Gl	A G u G	lu :	GAA Glu 245	GTA Val	738
20	GT. Va	A AT	T AG	GA TO	er va	C A	AT sn	TTC Phe	ACA Thr	A GAG	P As	T Go	er /	AAA Lys	ATC Ile	777
	AT: 11: 26:	= 11	A G: e Va	TA CA	G CT n Le	G A/ u Ly 26	/S	GAA Glu	CCT Pro	CTA Val	A GC	A A1 a I1 27	e A	TAF neF	TGT Cys	816
25	AC/ Thi	A AG	A CC g Pr 27	C AA O As	C AA n As	C AA n As	AT i	ACA Thr	AGA Arg 280	Lys	GG'	TA T	A C	TAT	CTA Leu 285	855
30	GG# Gly	CC Pr	A GG	G AG y Se	C AC. r Th. 29	r Ph	T :	PAT Pyr	ACA Thr	ACA Thi	GG/ G1 ₃ 295	, G1	A A u I	TA	ATA Ile	894
35	GGA Gly	GA As 30	5 II	A AG e Ar	A AA; g Ly:	A GC s Al	a 7	TAT Tyr 305	TGC Cys	AAG Lys	ATT	AG Se	r L	AA ys 10	GAA Glu	933
40	AAA Lys	TG	AA As	T AAG n Ası 319	1 Thi	r TT Le	A A u A	AGA Arg	CAG Gln	GTA Val 320	GTT Val	AA Ly	A A S I.	AA ' ys 1	TTA Leu	972
	AGA Arg 325	GA! Glu	CA.	A TTT	r GGC e Gly	330 330	n L	AA ys	ACA Th <i>r</i>	ATA Ile	ATT Ile	TT' Pho 33	⊋ A:	AT (CGA Arg	1011
45	TCC Ser	TCA	GG G1: 340	/ Gly	GAC Asp	CC/ Pro	A G	lu	ATT Ile 345	GTA Val	ATG Met	CAC	C AC	er F	TTT he	1050
50	AAC Asn	TGT Cys	GG/ Gly	GGG Gly	GAG Glu 355	Ph∈	T T	TC 1	TAC Tyr	TGT Cys	AAT Asn 360	AC#	A AC	CA C	AA 1n	1089
55	CTG Leu	TTT Phe 365	ASI	AGT Ser	ACT Thr	TGG) A:	AT / sn / 70	AAT Asn	ACT Thr	GAA Glu	GG G	AC Th 37	r A	AT : sn	1128
60	AGC Ser	ACT Thr	GAA Glu	GGA Gly 380	Asn	AGC Ser	Ti	CA A	le	ACA Thr 385	CTC Leu	CCA Pro	тс Су	C A s A	GA 1	1167
	ATA Ile 390	AAA Lys	CAA Gln	ATT Ile	ATA Ile	AAT Asn 395	Me	rg T et T	GG (CAG Gln	Glu	GTA Val 400	Gl	A A	AA 1 ys	206

	GCA	ACG	TAT	GCC	CCT	ccc	ATC	AGA	GGA	CGA	ATT	AGA	TGC	124	5
	Ala	Thr	Tyr 405	Ala	Pro	Pro	Ile	Arg 410	Gly	Arg	Ile	Arg	Cys 415		
5	ATA Ile	TCA Ser	AAT Asn	ATT Ile	ACA Thr 420	GGA Gly	CTG Leu	CTA Leu	TTA Leu	ACA Thr 425	AGA Arg	GAT Asp	GGT Gly	128	4
10	GGT Gly	AGG Arg 430	AAT Asn	GTC Val	ACA Thr	AAC Asn	AAT Asn 435	ACC Thr	GAA Glu	ACC Thr	TTC Phe	AGA Arg 440	CCT Pro	132	3
15	GGA Gly	GGA Gly	GGA Gly	GAC Asp 445	ATG Met	AGG Arg	GAC Asp	AAT Asn	TGG Trp 450	AGA Arg	AGT Ser	GAA Glu	TTA Leu	136	2
20	TAT Tyr 455	AAA Lys	TAT Tyr	AAA Lys	GTA Val	GTA Val 460	AAA Lys	GTT Val	GAA Glu	CCA Pro	TTA Leu 465	GGA Gly	ATA Ile	140	1
20	GCA Ala	CCC Pro	ACC Thr 470	AAG Lys	GCA Ala	AAG Lys	AGA Arg	AGA Arg 475	GTG Val	GTG Val	CAC His	AGA Arg	GAC Asp 480	1440	ס
25	AAA Lys	AGA Arg	GCA Ala	GCA Ala	CTA Leu 485	GGA Gly	GCC Ala	TTG Leu	TTC Phe	CTT Leu 490	GGG Gly	TTC Phe	TTA Leu	1479	Ð
30	GGA Gly	GCA Ala 495	TAA Xaa	AAG Lys	CTT Leu	CTA Leu 499	GA 1	1499						•	
35		L) SE (<i>F</i> (E	EQUEN () LE () TY	NCE (ENGTH (PE: OPOLO	HARA H: 49 Amir GY:	ACTEF 99 am 10 Ac Line	RISTI nino :id	CS: acid	is	10:18	3:				
40	Glu 1	Val	Pro	Val	Trp 5	Lys	Glu	Ala	Thr	Thr 10	Thr	Leu	Phe	Cys	Ala 15
	Ser	Asp	Ala	Lys	Ala 20	Tyr	Asp	Thr	Gly	Val 25	His	Asn	Val	Trp	Ala 30
45	Thr	His	Ala	Cys	Val 35	Pro	Thr	Asp	Pro	Asn 40	Pro	Gln	Glu	Ile	Glu 45
50	Leu	Val	Asn	Val	Thr 50		Asp			65	Trp	Lys	Asn	Lys	Met 60
	Val	Asp	Gln	Met	His 65	Glu	Asp	Ile	Ile	Ser 70	Leu	Trp	Asp	Glu	Ser 75
55	Leu	Lys	Pro	Cys	Val 80	Lys	Leu	Thr	Pro	Leu 85	Cys	Val	Thr	Leu	Asn 90
	Сув	Ser	Asp	Val	Asn 95	Asn	Ser	Thr	Asn	Pro 100	Asn	Asp	Thr	Asn	Thr 105
60	Asn	Ser	Thr	Asn	Thr 110	Thr	Ser	Ser	Thr	Pro 115	Thr	Ala	Thr	Thr	Ser 120
65	Ser	Glu	Glu	Lys	Met 125	Glu	Lys	Gly	Glu	11e 130	Lys	Asn	Cys	Ser	Phe 135

	Asn	Ile 1	hr Th	r Hi:	в Met О	Ly:	s As _l	p Ly	8 Ala 145	Glr	ı Lys	Glu	ту	r Ala 150
5	Leu 1	Phe 1	yr Ly	s Let 159	aA u	o Ile	e Vai	l Pro	11e	e Asp) yat) Asn	. Ası	n Ala 165
	Ser 1	Tyr A	rg Le	u Ile 170	e Ser	Cys	a Asr	Thi	ser 175	Asp	Ile	Thr	Glr	n Ala 180
10	Cys I	Pro L	ys Va	1 Thr 189	Phe	Glu	Pro	Ile	Pro 190	Ile	His	Tyr	Cys	3 Ala 195
15	Pro A	Ala G	ly Ph	e Ala 200	lle	Leu	Lys	Cys	205	Asp	Lys	Lys	Ph∈	210
	Gly T	Thr G	ly Pr	O Cys 215	Ser	Lys	. Val	Ser	Thr 220	Val	Gln	Cys	Thr	His 225
20	Gly I	le A	rg Pr	o Val 230	Val	Ser	Thr	Gln	Leu 235	Leu	Leu	Asn	Gly	Ser 240
	Leu A	la G	lu Gl	u Glu 245	Val	Val	Ile	Arg	Ser 250	Val	Asn	Phe	Thr	Asp 255
25	Asn A	la L	ys Il	260	Ile	Val	Gln	Leu	Lys 265	Glu	Pro	Val	Ala	Ile 270
30	Asn C			275					280					285
	Gly P	ro G	ly Sei	7hr 290	Phe	Tyr	Thr	Thr	Gly 295	Glu	Ile	Ile	Gly	Asp 300
35	Ile A	rg Ly	s Ala	305	Cys	Lys	Ile	Ser	Lys 310	Glu	Lys	Trp	Asn	Asn 315
	Thr L	eu Ar	g Glr	Val 320	Val	Lys	Lys	Leu	Arg 325	Glu	Gln	Phe	Gly	neA 330
40	Lys Ti			335					340					345
45	Val M∈	et Hi	s Ser	Phe 350	Asn	Cys	Gly	Gly	Glu 355	Phe	Phe	Tyr	Cys	Asn 360
	Thr Th	ır Gl	n Leu	Phe 365	Asn	Ser	Thr	Trp	Asn . 370	Asn	Thr	Glu (Gly	Thr 375
50	Asn Se	er Th	r Glu	Gly 380	Asn	Ser	Thr	Ile	Thr :	Leu	Pro (Cys /	Arg	Ile 390
	Lys Gl	in Il	e Ile	Asn 395	Met	Trp	Gln		Val (400	Gly 1	Lys 1	Ala ?		Tyr 405
55	Ala Pr	o Pr	o Ile	Arg 410	Gly .	Arg	Ile		Cys :	lle :	Ser A	Asn 1		Thr 420
60	Gly Le	u Le	ı Leu	Thr 425	Arg i	Asp	Gly		Arg <i>1</i> 430	Asn \	/al 1	Thr A		Asn 435
	Thr Gl	u Thi	r Phe	Arg 440	Pro (Gly (Gly (Asp N 445	det A	Arg A	sp A		Trṗ 450
65	Arg Se	r Glu	. Leu	Tyr 455	Lys 1	Tyr 1	Lys '		Val I 460	ys V	al G	lu P		Leu 465

	Gly Ile	Ala Pro	Thr Lys	Ala Lys	Arg Arg		l His	Arg Asp 480
5	Lys Arg	Ala Ala	Leu Gly 485	Ala Leu	Phe Let 490		e Leu	Gly Ala 495
	Xaa Lys	Leu Leu 499						
10	(1	RMATION EQUENCE A) LENGT B) TYPE: C) STRAN	CHARACTE H: 1499 Nucleic	RISTICS: base pai Acid				
15	1)	O) TOPOLO EQUENCE	OGY: Lin	ear	ID NO:1	9:		
20	GAG Glu 1	GTA CCT Val Pro	GTA TGG Val Trp 5	Lys Glu	GCA ACC Ala Thr	ACT ACT Th	T CTA r Leu	36
3.5	TTT TGT Phe Cys	GCA TCA Ala Ser 15	GAT GCT Asp Ala	AAA GCA Lys Ala 20	Tyr Asp	ACA GA Thr Gl	G GTG u Val 25	75
25	CAT AAT His Asn	GTT TGG Val Trp	GCC ACA Ala Thr 30	CAT GCC His Ala	TGT GTA Cys Val 35	Pro Th	A GAC r Asp	114
30	CCC AAC Pro Asn 40	CCA CAA Pro Gln	GAA ATA Glu Ile	GAA TTG Glu Leu 45	GTA AAT Val Asn	GTG AC Val Th 5	r Glu	153
35	GAT TTT Asp Phe	AAC ATG Asn Met 55	TGG AAA Trp Lys	AAT AAA Asn Lys	ATG GTA Met Val 60	GAC CA Asp Gl	G ATG n Met	192
40	CAT GAG His Glu 65	GAT ATA Asp Ile	ATC AGT Ile Ser 70	Leu Trp	GAT GAA Asp Glu	AGC CT Ser Le 75	A AAG u Lys	231
4.5	CCA TGT Pro Cys	GTA AAG Val Lys 80	TTA ACC Leu Thr	CCA CTT Pro Leu 85	Cys Val	ACT CT Thr Le	A AAC J Asn 90	270
45	TGC AGT Cys Ser	GAT GTG Asp Val	AAC AAT Asn Asn 95	TCC ACA Ser Thr	AAT CCT Asn Pro 100	Asn As	r ACT p Thr	309
50	AAT ACT Asn Thr 105	AAT TCC Asn Ser	ACT AAT Thr Asn	ACT ACT Thr Thr 110	TCC TCT Ser Ser	ACT CCT Thr Pro	o Thr	348
55	GCC ACT Ala Thr	ACT AGT Thr Ser 120	AGC GAG Ser Glu	GAA AAG Glu Lys	ATG GAG Met Glu 125	AAG GG Lys Gl	A GAA / Glu	387
60	ATA AAA Ile Lys 130	AAC TGC Asn Cys	TCT TTC Ser Phe 135	Asn Ile	ACC ACA Thr Thr	CAC ATO His Med 140	AAA Lys	426
	GAT AAG Asp Lys	GTA CAG Val Gln 145	AAA GAA Lys Glu	TAT GCA Tyr Ala 150	Leu Phe	TAT AA	CTT Leu 155	465
65								

-129-

	GA As	T AT p Il	A GT e Va	A CC l Pr	A ATA	e As	T GA	T AA p As	T AA n As	T AC n Th	r Se	C TA	T AG	G 504 9
5	TT Le	G AT u Il 17	e Se	T TG	T AAT	r AC	c Tc r Se 17	r Va	C AT	T AC	A CAC	G GC n Al 18	a Cy	T 543
10	CC. Pr	A AT	G GTG	G ACC	r Phe	GAC Glu	G CC	A AT	T CCC Pro 190	o Ile	A CAT	TA'	T TG: r Cys	r 582
15	GC6 A12 199	a Pr	G GC	r GG: a Gly	r TT1 / Phe	GCC Ala 200	a Ile	CT:	A AAC J Lys	TG1	AAA Lys 205	As	T AAC D Lys	621
20	AAC Lys	TTO Phe	AA1 Asr 210	ı Gly	ACA Thr	GGA Gly	CCP Pro	TG1 Cys 215	s Ser	AAC Lys	GTC Val	Se:	ACA Thr 220	
	GT# Val	CA/	TG1 Cys	ACA Thr	CAT His 225	GGA Gly	ATT	AGC Arç	CCA Pro	GTA Val 230	Val	TC# Ser	ACT Thr	699
25	CAA Gln	Leu 239	Leu	TTA Leu	AAT Asn	GGC	Ser 240	Leu	GCA Ala	GAA Glu	GAA Glu	GAA G1u 245	Val	738
30	GTA Val	ATT	AGA Arg	TCT Ser 250	GTC Val	AAT Asn	TTC Phe	ACA Thr	GAC Asp 255	AAT Asn	GCT Ala	AAA Lys	ATC Ile	777
35	ATA Ile 260	Ile	GTA Val	CAG Gln	CTG Leu	AAA Lys 265	GAA Glu	CCT Pro	GTA Val	GCA Ala	ATT Ile 270	AAT Asn	TGT Cys	816
40	ACA Thr	AGA Arg	CCC Pro 275	AAC Asn	AAC Asn	AAT Asn	ACA Thr	AGA Arg 280	AAA Lys	GGT Gly	ATA Ile	CAT His	CTA Leu 285	855
	GGA Gly	CCA Pro	GGG Gly	AGC Ser	ACA Thr 290	TTT Phe	TAT Tyr	ACA Thr	ACA Thr	GGA Gly 295	GAA Glu	ATA Ile	ATA Ile	894
45	GGA Gly	GAC Asp 300	ATA Ile	AGA Arg	AAA Lys	GCA Ala	TAT Tyr 305	TGC Cys	AAG Lys	ATT Ile	AGT Ser	AAA Lys 310	GAA Glu	933
50	AAA Lys	TGG Trp	Asn	AAC Asn 315	ACT Thr	Leu	Arg	Gln	GTA Val 320	Val	AAA Lys	AAA Lys	TTA Leu	972
55	AGA Arg 325	GAA Glu	CAA Gln	TTT Phe	GGG Gly	TAA neA 088	AAA Lys	ACA Thr	ATA Ile	ATT Ile	TTT Phe 335	AAT Asn	CGA Arg	1011
60	TCC Ser	TCA Ser	GGA Gly 340	GGG Gly	GAC Asp	CCA Pro	GAA Glu	ATT Ile 345	GTA Val	ATG Met	CAC His	AGT Ser	TTT Phe 350	1050
J.	AAC Asn	TGT Cys	GGA Gly	GGG	GAG Glu 355	TTT Phe	TTC Phe	TAC Tyr	Cys	AAT Asn '	ACA Thr	ACA Thr	CAA Gln	1089

	CTG Leu	TTT Phe 365	AAT Asn	AGT Ser	ACT Thr	TGG Trp	AAT Asn 370	AAT Asn	ACT Thr	GAA Glu	GGG Gly	ACA Thr 375	AAT Asn	112	8
5	AGC Ser	ACT Thr	GAA Glu	GGA Gly 380	TAA neA	AGC Ser	ACA Thr	ATC Ile	ACA Thr 385	CTC Leu	CCA Pro	TGC Cys	AGA Arg	116	7
10	Ile 390	Lys	Gln	Ile	Ile	AAT Asn 395	Met	Trp	Gln	Glu	Val 400	Gly	Lys		
15	Ala	Thr	Tyr 405	Ala	Pro	CCC Pro	Ile	Arg 410	Gly	Arg	Ile	Arg	Cys 415		
	ATA Ile	TCA Ser	AAT Asn	ATT	ACA Thr 420	GGA Gly	CTG Leu	CTA Leu	TTA Leu	ACA Thr 425	AGA Arg	GAT Asp	GGT Gly	1284	;
20	GGT Gly	AGG Arg 430	AAT Asn	GTC Val	ACA Thr	AAC Asn	AAT Asn 435	ACC Thr	GAN Xaa	NCC Xaa	TTC Phe	AGA Arg 440	CCT Pro	1323	3
25	GGA Gly	GGA Gly	GGA Gly	GAC Asp 445	ATG Met	AGG Arg	GAC Asp	AAT Asn	TGG Trp 450	AGA Arg	AGT Ser	GAA Glu	TTA Leu	1362	2
30	Tyr 455	Lys	Tyr	Lys	Val	GTA Val 460	Lys	Val	Glu	Pro	Leu 465	Gly	Ile		
35	Ala	Pro	Thr 470	Lys	Ala	AAG Lys	Arg	Arg 475	Val	Val	His	Arg	Asp 480		
	AAA Lys	AGA Arg	GCA Ala	GCA Ala	CTA Leu 485	GGA Gly	GCT Ala	TTG Leu	TTC Phe	CTT Leu 490	GGG	TTC Phe	TTA Leu	1479)
40	GGA Gly	GCA Ala 495	TAA Xaa	AAG Lys	CTT Leu	CTA Leu 499	GA 1	499							
45	(2) I (i	L) SE (<i>P</i> (E	QUEN A) LE B) TY	ICE C ENGTH (PE:	HARA : 49 Amir	CTER 9 am 10 Ac	ISTI ino id	cs:							
50	(x)	(C L) SE	O) TO	OPOLO	GY: ESCF	Line RIPTI	on:	SEQ	ID N	0:20):				
	Glu 1	Val	Pro	Val	Trp 5	Lys	Glu	Ala	Thr	Thr 10	Thr	Leu	Phe	Cys	Ala 15
55					20	Tyr				25					30
60					35	Pro				40					45
					50	Glu				55					60
65	Val	Asp	Gln	Met	His 65	Glu	Asp	Ile	Ile	Ser 70	Leu	Trp	Asp	Glu	Ser 75

	Leu	Lys	Pro	Су	s Val 80		: Le	u Thi	r Pro	o Lev 85		s Val	Th:	r Le	u Asn 90
5	Cys	Ser	. Ast	Va)	l Asr 95		s Se	Th:	c Ası	n Pro 100		n Asp	Th	r As	n Thr 105
	Asn	Ser	Thr	Ası	110		Ser	r Ser	Thi	Pro 115		Ala	Thi	r Th	r Ser 120
10	Ser	Glu	Glu	Lys	125	Glu	Lys	3 Gly	/ Glu	11e 130		Asn	Cys	s Se	Phe 135
15					140					145				_	7 Ala 150
					155					160					165
20					170					175					180
					185					190					Ala 195
25					200					205					Asn 210
30					Cys 215					220			-		225
	Gly				230					235					240
35	Leu				245					250					255
4.6	Asn				260					265					270
40	Asn				275					280					285
45	Gly				290					295					300
	Ile /				305					310					315
50	Thr I				320					325				-	330
5.5	Lys				335					340					345
55	Val 1				350					355				-	360
60	Thr T				365					370				-	375
	Asn S				380					385			-		390
65	Lys G	31n	Ile	Ile	Asn 395	Met '	Trp	Gln		Val (400	Gly 1	Lys i	Ala		Tyr 405

	Ala	Pro	Pro	Ile	Arg 410	Gly	Arg	Ile	Arg	Cys 415	Ile	Ser	Asn	Ile	Thr 420
5	Gly	Leu	Leu	Leu	Thr 425	Arg	Asp	Gly	Gly	Arg 430	Asn	Val	Thr	Asn	Asn 435
	Thr	Xaa	Xaa	Phe	Arg 440	Pro	Gly	Gly	Gly	Asp 445	Met	Arg	Asp	Asn	Trp 450
10	Arg	Ser	Glu	Leu	Tyr 455	Lys	Tyr	Lys	Val	Val 460	Lys	Val	Glu	Pro	Leu 465
1.5	Gly	Ile	Ala	Pro	Thr 470	Lys	Ala	Lys	Arg	Arg 475	Val	Val	His	Arg	Asp 480
15	Lys	Arg	Ala	Ala	Leu 485	Gly	Ala	Leu	Phe	Leu 490	Gly	Phe	Leu	Gly	Ala 495
20	Xaa	Lys	Leu	Leu 499											
25	į (<u>i</u>	i) SE (<i>P</i> (E (C	EQUEN A) LE B) TY C) SI O) TO	ICE C ENGTH PE: TRANK POLC	HARA H: 14 Nucl EDNE	SEQ 1 ACTER 175 t Leic ESS: Line RIPT	RIST: pase Acio Sino ear	CS: pain l	:s	10:2]	l:				
30	G	GTA Val 1	CCT Pro	GTG Val	TGG Trp	AAA Lys 5	GAA Glu	GCA Ala	AAC Asn	ACA Thr	ACT Thr 10	CTA Leu	TTT Phe	37	
35	TGT Cys	GCA Ala	TCA Ser 15	GAT Asp	GCT Ala	AAA Lys	GCA Ala	TAT Tyr 20	GAT Asp	AGA Arg	GAA Glu	GTA Val	CAT His 25	76	
40	AAT Asn	GTT Val	TGG Trp	GCA Ala	ACA Thr 30	CAT His	GCC Ala	TGT Cys	GTA Val	CCC Pro 35	ACA Thr	GAC Asp	CCC Pro	115	
45	AAC Asn	CCA Pro 40	CAA Gln	GAA Glu	ATA Ile	GTA Val	TTG Leu 45	GGA Gly	AAT Asn	GTG Val	ACA Thr	GAA Glu 50	AAT Asn	154	
45	TTT Phe	AAC Asn	ATG Met	TGG Trp 55	AAA Lys	AAT Asn	AAC Asn	ATG Met	GTA Val 60	GAA Glu	CAA Gln	ATG Met	CAT His	193	
50	GAG Glu 65	GAT Asp	ATA Ile	ATC Ile	AAT Asn	TTA Leu 70	TGG Trp	GAT Asp	CAA Gln	AGC Ser	TTA Leu 75	AAG Lys	CCA Pro	232	
55	TGT Cys	GTA Val	AAG Lys 80	TTA Leu	ACT Thr	CCA Pro	CTC Leu	TGT Cys 85	GTT Val	ACT Thr	TTA Leu	AAG Lys	TGC Cys 90	271	
60	AAG Lys	GAT Asp	CTG Leu	GAG Glu	AGG Arg 95	AAT Asn	ACT Thr	ACC Thr	TAT Tyr	AAT Asn 100	AGC Ser	ACT Thr	ATT Ile	310	
	ACC Thr	AAT Asn 105	AAT Asn	AGT Ser	AGT Ser	TTG Leu	GAG Glu 110	GGA Gly	CTA Leu	AGA Arg	GAA Glu	CAA Gln 115	ATG Met	349	
65															

-133-

	ACA Thr	AAC A n	TGC Cys	TCT Ser 120	Phe	AAC Asn	ATC Ile	ACC Thi	2 ACA Thi 129	Ser	ATA Ile	AGA Arg	GAT Asp	388
5	AAG Lys 130	Val	CAG Gln	AAA Lys	GAA Glu	TAT Tyr 135	GCA Ala	CT1	TTC Let	TAT Tyr	Lys 140	Leu	GAT Asp	427
10	GTA Val	GTA Val	CCA Pro 145	ATA Ile	GAA Glu	GAA Glu	GAT Asp	GAC Asp 150	Asn	ACT Thr	AGC Ser	TAT Tyr	AGA Arg 155	
15	TTG Leu	ATA Ile	AGT Ser	TGT Cys	AAC Asn 160	ACC Thr	TCA Ser	GTC Val	: ATT Ile	ACA Thr 165	CAG Gln	GCT Ala	TGT Cys	505
20	CCA Pro	AAG Lys 170	ACA Thr	TCC Ser	TTT Phe	GAG Glu	CCA Pro 175	ATT	CCC Pro	ATA Ile	CAT His	TAT Tyr 180	TGT Cys	544
	GCC Ala	CCG Pro	GCT Ala	GGT Gly 185	TTT Phe	GCG Ala	ATT Ile	CTA Leu	ANG Lys 190	TGT Cys	AAT Asn	GAT Asp	AAG Lys	583
25	AAG Lys 195	TTC Phe	AAT Asn	GGA Gly	ACA Thr	GGA Gly 200	CCA Pro	TGT Cys	AAA Lys	AAT Asn	GTC Val 205	AGC Ser	ACA Thr	622
30	GTA Val	CAA Gln	TGT Cys 210	ACA Thr	CAT His	GGA Gly	ATT Ile	AGG Arg 215	CCA Pro	GTA Val	GTA Val	TCA Ser	ACT Thr 220	661
35	CAA Gln	CTG Leu	TTG Leu	TTA Leu	AAT Asn 225	GGC Gly	AGT Ser	CTA Leu	GCA Ala	GAA Glu 230	GAA Glu	GAG Glu	GTA Val	700
40	GTA Val	ATC Ile 235	AGA Arg	TCT Ser	GCC Ala	AAT Asn	TTC Phe 240	ACA Thr	GAC Asp	AAT Asn	GCT Ala	AAA Lys 245	ACC Thr	739
40	ATA Ile	ATA Ile	GTA Val	CAT His 250	CTA Leu	AAT Asn	GAA Glu	ACT Thr	GTA Val 255	AAA Lys	ATT Ile	AAT Asn	TGT Cys	778
45	ACA Thr 260	AGA Arg	CTT Leu	GGC Gly	AAC Asn	AAT Asn 265	ACA Thr	AGA Arg	AAA Lys	AGT Ser	ATA Ile 270	AAT Asn	ATA Ile	817
50	GGA Gly	CCA Pro	GGG Gly 275	AGA Arg	GTA Val	CTC Leu	TAT Tyr	GCA Ala 280	ACA Thr	GGA Gly	GAA Glu	ATA Ile	ATA Ile 285	856
55	GGA Gly	GAC Asp	ATA Ile	Arg	CAA Gln 290	GCA Ala	CAT His	TGT Cys	AAC Asn	ATT Ile 295	AGT Ser	AGA Arg	GCA : Ala	895
60	CAA Gln	TGG Trp 300	AAT Asn	AAG Lys	ACT Thr	Leu (GAA Glu 305	AAG Lys	CTA Val	GTT Val	Asp	AAA Lys 310	TTA S Leu	934
60	AGA Arg	AAA Lys	Gln	TTT Phe	GGG (GAT A	AAT Asn	ACA Thr	ACA Thr 320	ATA Ile	GCT Ala	TTT Phe	AAT 9 Asn	973

						GAC Asp 330								101	2
5						GAA Glu								105	1
10	CAA Gln	CTG Leu	TTT Phe	AAT Asn	AGT Ser 355	ACT Thr	TGG Trp	TAA neA	AAT Asn	ACT Thr 360	TGG Trp	AAG Lys	GAT Asp	109	0
15	CCT Pro	AAC Asn 365	AGG Arg	AGT Ser	GAC Asp	AAT Asn	ATC Ile 370	ACA Thr	CTC Leu	CCA Pro	TGC Cys	AGA Arg 375	ATA Ile	1129	€
	AAA Lys	CAA Gln	ATT Ile	ATA Ile 380	AAC Asn	ATG Met	TGG Trp	CAG Gln	GAA Glu 385	GTA Val	GGA Gly	AAA Lys	GCA Ala	1168	3
20	ATG Met 390	TAC Tyr	GCC Ala	CCT Pro	CCC Pro	ATC Ile 395	AGA Arg	GGG Gly	GAA Glu	ATT Ile	AGA Arg 400	TGT Cys	TCA Ser	1207	7
25	TCA Ser	AAT Asn	ATC Ile 405	ACA Thr	GGG Gly	CTG Leu	CTA Leu	CTA Leu 410	ACA Thr	AGA Arg	GAT Asp	GGT Gly	GGT Gly 415	1246	5
30	AAT Asn	GAC Asp	GAT Asp	GGT Gly	AAT Asn 420	GAC Asp	ACG Thr	ACC Thr	ACA Thr	AAC Asn 425	AGG Arg	ACC Thr	GAG Glu	1285	•
35	ATC Ile	TTC Phe 430	AGA Arg	CCT Pro	GGA Gly	GGA Gly	GGA Gly 435	GAT Asp	ATG Met	AGG Arg	GAC Asp	AAT Asn 440	TGG Trp	1324	,
4.0	AGA Arg	AGT Ser	GAA Glu	TTA Leu 445	TAT Tyr	AGA Arg	TAT Tyr	AAA Lys	GTA Val 450	GTA Val	AAA Lys	ATT Ile	GAA Glu	1363	,
40	CCA Pro 455	TTA Leu	GGA Gly	ATA Ile	GCA Ala	CCC Pro 460	acc Thr	AGG Arg	GCA Ala	AAG Lys	AGA Arg 465	AGA Arg	GTG Val	1402	!
45	GTG Val	CAG Gln	AGA Arg 470	GAA Glu	AAA Lys	AGA Arg	GCA Ala	GTA Val 475	GGA Gly	CTA Leu	GGA Gly	GCT Ala	TTG Leu 480	1441	
50		CTT Leu		т тс	TTAG	GAG	CATA	AAGC	TT C	TAGA	147	5			
55	·	.) SE (A (B (D	QUEN () LE () TY	ICE C INGTH IPE: IPOLO	HARA : 49 Amin GY:	EQ I CTER 1 am o Ac Line IPTI	ISTI ino id ar	CS: acid	s	0.22					
60	-					Glu						Phe	Суз	Ala	Ser 15
65	Asp	Ala	Lys	Ala	Tyr 20	Asp	Arg	Glu	Val	His 25	Asn	Val	Trp	Ala	Thr 30

-135-

	His	Ala	Cys	s Val	l Pro 39	Th:	c As	p Pr	o As	n Pro 40	Gln	Glu	Ile	Va	l Leu 45
5	Gly	Asn	Val	Thr	G10 50	Asr	Ph	e As	n Me	t Trp 55	Lys	Asn	Asn	Met	t Val
	Glu	Gln	Met	His	65 65	Asp) Il	e Il	e As	n Leu 70	Trp	Asp	Gln	Ser	Leu 75
10	Lys	Pro	Cys	Val	Lys 80	Leu	Th	Pro) Le	u Cys 85	Val	Thr	Leu	Lys	Cys 90
15	Lys	Asp	Leu	Glu	Arg 95	Asn	Thi	Thi	ту	r Asn 100	Ser	Thr	Ile	Thr	Asn 105
15	Asn	Ser	Ser	Leu	Glu 110	Gly	Leu	Arç	g Glu	3 Gln 115	Met	Thr	Asn	Cys	Ser 120
20	Phe	Asn	Ile	Thr	Thr 125	Ser	Ile	Arg	, Asp	Lys 130	Val	Gln	Lys	Glu	Tyr 135
	Ala	Leu	Leu	Tyr	Lys 140	Leu	Asp	Val	Val	Pro 145	Ile	Glu	Glu	Asp	Asp 150
25	Asn	Thr	Ser	Tyr	Arg 155	Leu	Ile	Ser	Cys	Asn 160	Thr :	Ser	Val	Ile	Thr 165
30	Gln	Ala	Cys	Pro	Lys 170	Thr	Ser	Phe	Glu	Pro 175	Ile 1	Pro	Ile	His	Tyr 180
30	Cys	Ala	Pro	Ala	Gly 185	Phe	Ala	Ile	Leu	Lys (Cys A	Asn i	Asp 1	Ĺys	Lys 195
35	Phe .	Asn	Gly	Thr	Gly 200	Pro	Cys	Lys	Asn	Val 3 205	Ser 1	hr 1	Val (Sln	Cys 210
	Thr I	His	Gly	Ile	Arg 215	Pro	Val	Val	Ser	Thr 0	Gln L	eu I	Ceu I		Asn 225
40	Gly s	Ser :	Leu .	Ala	Glu 230	Glu	Glu	Val	Val	Ile A 235	Arg S	er A	Nla A		Phe 240
45	Thr A	Asp A	Asn i	Ala	Lys 245	Thr	Ile	Ile	Val	His L 250	.eu A	sn G	lu T		Val 255
43	Lys I	le /	Asn (Cys	Thr 260	Arg 1	Leu	Gly	Asn	Asn T 265	hr A	rg L	ys S		Ile 270
50	Asn I	le o	Sly I	Pro (Gly 1 275	Arg '	Val	Leu	Tyr	Ala T 280	hr G	ly G	lu I		Ile 285
	Gly A	sp I	le A	Arg (31n / 290	Ala F	dis	Cys	Asn	Ile S 295	er Aı	rg A	la G		Trp 300
55	Asn L	ys I	hr I	eu (31u I	Lys \	/al '	Val.	Asp	Lys L 310	eu Ai	g L	ys G		Phe 315
60	Gly A	sp A	sn T	hr 1	hr 1 320	le A	vla 1	Phe .	Asn	Arg So	er Se	er G	ly G	ly A	
00	Pro G	lu I	le V	al M	let H	lis T	hr 1	Phe A	Asn (Cys G:	ly Gl	y G	lu Ph	ne P	
65	Tyr C	ys A	sn T	hr T	hr G 50	ln L	eu 1	Phe A	Asn :	Ser Th 355	ır Tr	p As	sn As	n T	

	Trp	Lys	Asp	Pro	As n 365	Arg	Ser	Asp	Asn	Ile 370		Leu	Pro	Cys	Arg 375
5	Ile	Lys	Gln	Ile	11e 380	Asn	Met	Trp	Gln	Glu 385	Val	Gly	Lys	Ala	Met 390
	Tyr	Ala	Pro	Pro	Ile 395	Arg	Gly	Glu	Ile	Arg 400	Cys	Ser	Ser	Asn	11e 405
10	Thr	Gly	Leu	Leu	Leu 410	Thr	Arg	Asp	Gly	Gly 415	Asn	Asp	Asp	Gly	Asn 420
15	Asp	Thr	Thr	Thr	Asn 425	Arg	Thr	Glu	Ile	Phe 430	Arg	Pro	Gly	Gly	Gly 435
13	Asp	Met	Arg	Asp	Asn 440	Trp	Arg	Ser	Glu	Leu 445	Tyr	Arg	Tyr	Lys	Val 450
20	Val	Lys	Ile	Glu	Pro 455	Leu	Gly	Ile	Ala	Pro 460	Thr	Arg	Ala	Lys	Arg 465
	λrg	Val	Val	Gln	Arg 470	Glu	Lys	Arg	Ala	Val 475	Gly	Leu	Gly	Ala	Leu 480
25	Phe	Leu	Gly	Phe	Leu 485	Gly	Ala	Leu	Phe	Leu 490	Gly 491				
30	(2) I	i) SE (7 (E	EQUEN A) LE B) TY C) ST	NCE C ENGTH (PE: TRAND	HARA I: 14 Nucl	CTER 175 b .eic SSS:	RISTI ase Acid Sing	CS: pair i							
35	(x)	L) SE	O) TO	POLC	GY: ESCF	Line	ON:	SEQ	ID N	10:23	3:				
33	G	GTA Val 1	CCT Pro	GTG Val	TGG Trp	AAA Lys 5	GAA Glu	GCA Ala	AAC Asn	ACA Thr	ACT Thr 10	CTA Leu	TTT Phe	37	
40	TGT Cys	GCA Ala	TCA Ser 15	GAT Asp	GCT Ala	AAA Lys	GCA Ala	TAT Tyr 20	GAT Asp	AGA Arg	GAA Glu	GTA Val	CAT His 25	76	
45	AAT Asn	GTT Val	TGG Trp	GCA Ala	ACA Thr 30	CAT His	GCC Ala	TGT Cys	GTA Val	CCC Pro 35	ACA Thr	GAC Asp	CCC Pro	115	
50	AAC Asn	CCA Pro 40	CAA Gln	GAA Glu	ATA Ile	GTA Val	TTG Leu 45	GGA Gly	AAT Asn	GTG Val	ACA Thr	GAA Glu 50	TAA Asn	154	
	TTT Phe	AAC Asn	ATG Met	TGG Trp 55	AAA Lys	TAA Asn	AAC Asn	ATG Met	GTA Val 60	GAA Glu	CAA Gln	ATG Met	CAT His	193	
55	GAG Glu 65	GAT Asp	ATA Ile	ATC Ile	AAT Asn	TTA Leu 70	TGG Trp	GAT Asp	CAA Gln	AGC Ser	TTA Leu 75	AAG Lys	CCA Pro	232	
60	TGT Cys	GTA Val	AAG Lys 80	TTA Leu	ACT Thr	CCA Pro	CTC Leu	TGT Cys 85	GTT Val	ACT Thr	TTA Leu	AAG Lys	TGC Cys 90	271	

	AAG Lys	GAT Asp	CTG Leu	GAG Glu	AGG Arg) As	T AC' n Th	r AC	C TA	T AA' r Ası 100	n Sei	C AC	T AT	T 310
5	ACC Thr	AAT Asn 105	AAT Asn	AGT Ser	AG1 Ser	TTC Let	G GAG Glu 110	Gly	A CT	A AGA	A GAZ Glu	CA.	n Mei	3 349 t
10	ACA Thr	AAC Asn	TGC Cys	TCT Ser 120	TTC Phe	AAC Ası	TATO	C ACC	C ACA	Ser	T ATA	AG Ar	A GA1 g Ası	r 388 >
15	AAG Lys 130	GTG Val	CAG Gln	AAA Lys	GAA Glu	TAT Tyr 135	. Ala	CTT Leu	TTC Lev	TAT Tyr	Lys 140	Let	r GA1	427
20	GTA Val	GTA Val	CCA Pro 145	ATA Ile	GAA Glu	GAA Glu	GAT Asp	GAC Asp 150) Asn	ACT Thr	AGC Ser	TAT	AGA Arg 155	
	TTG Leu	ATA Ile	AGT Ser	TGT Cys	AAC Asn 160	Thr	TCA Ser	GTC Val	ATT Ile	ACA Thr 165	Gln	GCT Ala	TGT Cys	505
25	CCA Pro	AAG Lys 170	ACA Thr	TCC Ser	TTT Phe	GAG Glu	CCA Pro 175	ATT Ile	CCC Pro	ATA Ile	CAT His	TAT Tyr 180	Cys	544
30	GCC Ala	CCG Pro	GCT Ala	GGT Gly 185	TTT Phe	GCG Ala	ATT Ile	CTA Leu	AAG Lys 190	TGT Cys	AAT Asn	GAT Asp	AAG Lys	583
35	AAG Lys 195	TTC Phe	AAT Asn	GGA Gly	ACA Thr	GGA Gly 200	CCA Pro	TGT Cys	AAA Lys	AAT Asn	GTC Val 205	AGC Ser	ACA Thr	622
40	GTA Val	Gln	TGT Cys 210	ACA Thr	CAT His	GGA Gly	ATT	AGG Arg 215	CCA Pro	GTA Val	GTA Val	TCA Ser	ACT Thr 220	661
	CAA (Gln)	CTG ' Leu !	TTG Leu	Leu	AAT Asn 225	GGC Gly	AGT Ser	CTA Leu	GCA Ala	GAA Glu 230	GAA Glu	GAG Glu	GTA Val	700
45	GTA /	ATC I Ile I 235	AGA '	TCT Ser	GCC Ala	AAT Asn	TTC Phe 240	ACA Thr	GAC Asp	AAT Asn	GCT Ala	AAA Lys 245	ACC Thr	739
50	ATA I	ATA (Ile \	/al	CAT His 250	CTA Leu	AAT Asn	GAA Glu	ACT Thr	GTA Val 255	AAA Lys	ATT Ile	AAT Asn	TGT Cys	778
55	ACA A Thr A 260	AGA (Arg I	CTT (Leu (GGC . Gly	AAC Asn	AAT Asn 265	ACA Thr	AGA Arg	AAA Lys	AGT Ser	ATA Ile 270	TAA nea	ATA Ile	817
60	GGA G	Pro C	GG / Sly / 275	AGA (Arg '	GTA Val	CTC Leu	TAT Tyr	GCA Ala 280	ACA Thr	GGA Gly	GAA Glu	ATA Ile	ATA Ile 285	856
	GGA G	GAC A	TA A	arg (CAA Gln 290	GCA Ala	CAT His	TGT Cys	AAC Asn	ATT Ile 295	AGT . Ser .	AGA Arg	GCA Ala	895

	CAA Gln	TGG Trp 300	AAT Asn	AAG Lys	ACT Thr	TTA Leu	GAA Glu 305	AAG Lys	GTA Val	GTT Val	GAC Asp	AAG Lys 310	Leu	934
5	AGA Arg	AAA Lys	CAA Gln	TTT Phe 315	Gly	GAT Asp	AAT Asn	ACA Thr	ACA Thr 320	Ile	GCT Ala	TTT Phe	AAT Asn	973
10	CGA Arg 325	TCC Ser	TCA Ser	GGA Gly	GGG Gly	GAC Asp 330	CCA Pro	GAA Glu	ATT Ile	GTA Val	ATG Met 335	CAC His	ACT Thr	1012
15	TTT Phe	AAT Asn	TGT Cys 340	GGA Gly	GGG Gly	GAA Glu	TTT Phe	TTC Phe 345	Tyr	TGT Cys	AAT Asn	ACA Thr	ACA Thr 350	1051
							TGG Trp							1090
20	CCT Pro	AAC Asn 365	AGG Arg	AGT Ser	GAC Asp	AAT Asn	ATC Ile 370	ACA Thr	CTC Leu	CCA Pro	TGC Cys	AGA Arg 375	ATA Ile	1129
25	AAA Lys	CAA Gln	ATT Ile	ATA Ile 380	AAC Asn	ATG Met	TGG Trp	CAG Gln	GAA Glu 385	GTA Val	GGA Gly	AAA Lys	GCA Ala	1168
30	ATG Met 390	TAC Tyr	GCC Ala	CCT Pro	CCC Pro	ATC Ile 395	AGA Arg	GGG Gly	GAA Glu	ATT Ile	AGA Arg 400	TGT Cys	TCA Ser	1207
35	TCA Ser	AAT Asn	ATC Ile 405	ACA Thr	GGG Gly	CTG Leu	CTA Leu	CTA Leu 410	ACA Thr	AGA Arg	GAT Asp	GGT Gly	GGT Gly 415	1246
	AAT Asn	GAC Asp	GAT Asp	GGT Gly	AAT Asn 420	GAC Asp	ACG Thr	ACC Thr	ACA Thr	AAC Asn 425	AGG Arg	ACC Thr	GAG Glu	1285
40	ATC Ile	TTC Phe 430	AGA Arg	CCT Pro	GGA Gly	GGA Gly	GGA Gly 435	GAT Asp	ATG Met	AGG Arg	GAC Asp	AAT Asn 440	TGG Trp	1324
45	AGA Arg	AGT Ser	GAA Glu	TTA Leu 445	TAT Tyr	AGA Arg	TAT Tyr	AAA Lys	GTA Val 450	GTA Val	AAA Lys	ATT Ile	GAA Glu	1363
50	CCA Pro 455	TTA Leu	GGA Gly	ATA Ile	GCA Ala	CCC Pro 460	ACC Thr	AGG Arg	GCA Ala	AAG Lys	AGA Arg 465	AGA Arg	GTG Val	1402
55	GTG Val	CAG Gln	AGA Arg 470	GAA Glu	AAA Lys	AGA Arg	GCA Ala	GTA Val 475	GGA Gly	CTA Leu	GGA Gly	GCT Ala	TTG Leu 480	1441
							GCA Ala				Xaa	A 14	75	
60	(2) I (i) SE	QUEN	CE C	HARA 1: 49	CTER 1 am	ISTI ino	cs:						
65		(B) TY	PE:	Amin GY:	o Ac Line	id							

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:24:

5	Val 1	Pro	o Val	Tr	p Lys	Glu	a Ala	a Ası	Th:	r Thi		ı Phe	≘ Су	s Al	a Ser 15
-	Asp	Ala	Lys	s Ala	20 20	Asp	Arç	g Glu	ı Va	l His		ı Va	l Tr	p Al	a Thr 30
10	His	Ala	Cys	va]	Pro 35	The	. Ast	Pro) Asr	Pro 40		Glu	: Il	e Va	1 Leu 45
	Gly	Asn	val	Thr	Glu 50	Asn	Phe	: Asn	Met	Trp 55		Asr	ASI	n Me	t Val 60
15	Glu	Gln	Met	His	Glu 65	Asp	Ile	lle	Asn	Leu 70	Trp	Asp	Gli	ı Se	r Leu 75
20					80					85					90
					95					100					105
25					110					115					Ser 120
					Thr 125					130					135
30					Lys 140					145					150
35					Arg 155					160					165
					Lys 170					175					180
40 .	Cys				185					190					195
	Phe				200					205					210
45	Thr				215					220					225
50	Gly :				230					235					240
	Thr				245					250					255
55	Lys :				260					265					270
	Asn :				275					280					285
60	Gly 7				290					295					300 .
65	Asn I	ys	Thr	Leu	Glu : 305	Lys	Val '	Val .		Lys 1 310	Leu i	Arg 1	Ĺув		Phe 315

	Gly	Asp	Asn	Thr	Thr 320	Ile	Ala	Phe	Asn	Arg 325	Ser	Ser	Gly	Gly	Asp 330
5	Pro	Glu	Ile	Val	Met 335	His	Thr	Phe	Asn	Cys 340	Gly	Gly	Glu	Phe	Phe 345
	Tyr	CAa	Asn	Thr	Thr 350	Gln	Leu	Phe	Asn	Ser 355	Thr	Trp	Asn	Asn	Thr 360
10	Trp	Lys	Asp	Pro	Asn 365	Arg	Ser	Asp	Asn	11e 370	Thr	Leu	Pro	Cys	Arg 375
15				Ile	380					385					390
13	Tyr	Ala	Pro	Pro	Ile 395	Arg	Gly	Glu	Ile	Arg 400	Cys	Ser	Ser	Asn	Ile 405
20	Thr	Gly	Leu	Leu	Leu 410	Thr	Arg	Asp	Gly	Gly 415	Asn	Asp	Asp	Gly	Asn 420
	Asp	Thr	Thr	Thr	Asn 425	Arg	Thr	Glu	Ile	Phe 430	Arg	Pro	Gly	Gly	Gly 435
25	Asp	Met	Arg	Asp	Asn 440	Trp	Arg	Ser	Glu	Leu 445	Tyr	Arg	Tyr	Lys	Val 450
20	Val	Lys	Ile	Glu	Pro 455	Leu	Gly	Ile	Ala	Pro 460	Thr	Arg	Ala	Lys	Arg 465
30	Arg	Val	Val	Gln	Arg 470	Glu	Lys	Arg	Ala	Val 475	Gly	Leu	Gly	Ala	Leu 480
35	Phe	Leu	Gly	Phe	Leu 485	Gly	Ala	Xaa	Ser	Phe 490	Xaa 491				
	(2)	i) S1	13UQ3	ON INCE O	HAR	CTE	RIST	cs:							
40		(1	B) T' C) S' C) T(PE: PRANI	Nucl EDNE OGY:	leic ESS: Line	Acid Sind	i gle							
	(x	i) Šī	EQUE	NCE I	DESC	RIPT	ON:	SEQ	ID N	10:25	5:				
45		CTC Leu 1	GAG Glu	GTA Val	CCT Pro	GTG Val 5	TGG Trp	AAA Lys	GAA Glu	GCA Ala	ACC Thr 10	ACC Thr	ACT Thr	36	
50				GCA Ala										75	
55	GCA Ala	CAT His	AAT Asn	GTT Val	TGG Trp 30	GCC Ala	ACA Thr	CAT His	GCC Ala	TGT Cys 35	GTA Val	CCC Pro	ACA Thr	114	
	GAC Asp	CCC Pro 40	Asn	CCA Pro	CAA Gln	GAA Glu	GTA Val 45	GAA Glu	TTG Leu	GAA Glu	AAT Asn	GTG Val 50	ACA Thr	153	
60	GAA Glu	AAT	ттт	AAC Asn 55	ATG Met	TGG Trp	AAA	AAT Asn	AAC Asn 60	ATG Met	GTA Val	GAA	CAG Gln	192	

	ATO Met 65	: His	GGG Gly	GA]	ATA Ile	ATT	e Se	r TT	A TGC	G GAT	CAI Gli 7!	າ Se	C CT	A 231
5	AAC Lys	CCA Pro	TGT Cys 80	Val	AAA Lys	TT!	A ACC	C CC/ Pro 85	Lev	TG1	r GT1	r ACC	TT/ Let 90	
10	AAT Asn	TGC Cys	ACT Thr	GAC Asp	CCA Pro 95	Asr	GT7 Val	T AC1	TAA 1 Asn	Ser 100	Glu	AG/	A ACC	309
15	ATA Ile	GAG Glu 105	Gly	GGA Gly	GAA Glu	ATA	AAA Lys 110	: Asn	TGC Cys	TCT Ser	TTC Phe	AAT Asr	ılle	348
20	ACC Thr	ACA Thr	AAC Asn	ATA Ile 120	Arg	GAT Asp	AGG Arg	TTT Phe	CAG Gln 125	Lys	GAA Glu	TAT	GCA Ala	387
20	CTT Leu 130	Phe	TAT Tyr	AAA Lys	CTT Leu	GAT Asp 135	Val	ATA	CCA Pro	TTA Leu	GGT Gly 140	Asn	GAT Asp	426
25	AAT Asn	ACT Thr	AGC Ser 145	TAT Tyr	AGG Arg	TTG Leu	ATA Ile	AGT Ser 150	TGT Cys	AAC Asn	ACC Thr	TCA Ser	GTC Val 155	465
30	ATT Ile	ACA Thr	CAG Gln	GCC Ala	TGT Cys 160	CCA Pro	AAG Lys	GTA Val	TCC Ser	TTT Phe 165	GAG Glu	CCA Pro	ATT	504
35	CCC Pro	ATA Ile 170	CAT His	TAT Tyr	TGT Cys	GCC Ala	CCG Pro 175	GCT Ala	GGT Gly	TTT Phe	CCG Ala	ATT Ile 180	CTA Leu	543
40	AAG Lys	TGT Cys	AAA Lys	GAT Asp 185	AAG Lys	AAG Lys	TTC Phe	AAT Asn	GGA Gly 190	ACA Thr	GGA Gly	CCA Pro	TGT Cys	582
40	ACA Thr 195	AAT Asn	GTC Val	AGC Ser	ACA Thr	GTA Val 200	CAA Gln	TGT Cys	ACA Thr	CAT His	GGA Gly 205	ATT Ile	AAG Lys	621
45	CCA Pro	GTA Val	GTA Val 210	TCA Ser	ACT Thr	CAA Gln	CTG Leu	TTG Leu 215	TTA Leu	AAT Asn	GGC Gly	AGT Ser	CTA Leu 220	660
50	GCA Ala	GAA Glu	GAA Glu	Asp	ATA Ile 225	GTA Val	ATT Ile	AGA Arg	Ser	GCC Ala 230	AAT Asn	CTC Leu	ACA Thr	699
55	GAC Asp	AAT Asn 235	GCT Ala	AAA Lys	AAC Asn	ATA Ile	ATA Ile 240	GTA Val	CAG Gln	CTG Leu	TAA Asn	GAA Glu 245	TCT Ser	738
50	GTA Val	ACA Thr	Met .	AAT Asn 250	TGT Cys	ACA Thr	AGA Arg	CCC Pro	AAC Asn 255	AAC . Asn .	AAT Asn	ACA Thr	ATG Met	777
30	AAA Lys 260	AGT Ser	ATA (CAT His	Ile (GGA Gly 265	CCA Pro	GGC Gly	ACA (Ala :	TTT Phe	TAT Tyr	GCA Ala	816

	ACA Thr	GGA Gly	AAC Asn 275	ATA Ile	ATA Ile	GGA Gly	GAT Asp	ATA Ile 280	AGA Arg	CAA Gln	GCA Ala	CAT His	TGT Cys 285	855
5	AAC Asn	ATT Ile	AGT Ser	GGA Gly	ACA Thr 290	AAA Lys	TGG Trp	AAT Asn	GAC Asp	ACT Thr 295	TTG Leu	AAA Lys	AAG Lys	894
10	ATA Ile	GCT Ala 300	ATA Ile	AAA Lys	TTA Leu	AGA Arg	GAA Glu 305	CAA Gln	TTT Phe	AAT Asn	AAG Lys	ACA Thr 310	ATA Ile	933
15	Val	Phe	Asn	Gln 315	Ser	Ser	Gly	Gly	Asp 320	Pro	Glu	Ile		
20	ACG Thr 325	CTC Leu	AGT Ser	TTT Phe	AAT Asn	TGT Cys 330	GGA Gly	GGG Gly	GAA Glu	TTT Phe	TTC Phe 335	TAC Tyr	TGT Cys	1011
20	AAT Asn	TCA Ser	ACA Thr 340	CAA Gln	CTG Leu	TTT Phe	AAT Asn	AGT Ser 345	ACT Thr	TGG Trp	AAT Asn	AGT Ser	ACT Thr 350	1050
25	GGG Gly	TCA Ser	AAT Asn	AAC Asn	ACT Thr 355	AAA Lys	GGA Gly	AAT Asn	GAC Asp	ACA Thr 360	ATC Ile	ACA Thr	CTC Leu	1089
30 .	CCA Pro	TGC Cys 365	AGA Arg	ATA Ile	AGA Arg	CAA Gln	ATT Ile 370	ATA Ile	AAC Asn	ATG Met	TGG Trp	CAG Gln 375	AAA Lys	1128
35	ATA Ile	GGA Gly	AAA Lys	GCA Ala 380	ATG Met	TAT Tyr	GCC Ala	CCT Pro	CCC Pro 385	ATC Ile	AAA Lys	GGG Gly	CAA Gln	1167
4.0	ATT Ile 390	AGA Arg	TGT Cys	TCA Ser	TCA Ser	AAT Asn 395	ATT Ile	ACA Thr	GGG Gly	CTA Leu	ATA Ile 400	TTA Leu	ACA Thr	1206
40	AGA Arg	GAT Asp	GGT Gly 405	GGT Gly	AAC Asn	AAC Asn	AAC Asn	ATG Met 410	AGC Ser	AAG Lys	ACC Thr	ACC Thr	GAG Glu 415	1245
45	ACC Thr	TTC Phe	AGA Arg	CCT Pro	GGA Gly 420	GGA Gly	GGA Gly	GAT Asp	ATG Met	AGG Arg 425	GAC Asp	AAT Asn	TGG Trp	1284
50	AGA Arg	AGT Ser 430	GAA Glu	TTA Leu	TAT Tyr	AAA Lys	TAT Tyr 435	AAA Lys	GTA Val	GTA Val	AAA Lys	ATT Ile 440	GAA Glu	1323
55	CCA Pro	TTA Leu	GGA Gly	GTA Val 445	GCA Ala	CCC Pro	ACC Thr	AGG Arg	GCA Ala 450	AAG Lys	AGA Arg	AGA Arg	GTG Val	1362
	GTG Val 455	CAG Gln	AGA Arg	GAA Glu	AAA Lys	AGA Arg 460	GCA Ala	GTG Val	GGA Gly	ATA Ile	GGA Gly 465	GCT Ala	GTG Val	1401
60	TTC Phe	CTT Leu	GGG Gly 470	TTC Phe	TTG Leu	GGA Gly	GCA Ala	TAA Xaa 475	AGC Ser	TTC Phe	TAG Xaa 478	A 14	35	

5		(i) :	SEQUI (A) 1 (B) 1 (D) 1	ENCE LENG? LYPE: LOPOI	FOR CHAP TH: 4 : Ami LOGY: DESC	RACTE 178 a ino A : Lir	RIST mind wid wear	rics o ac	: ids	NO: 2	26:				
10	Le	u Glu 1	u Val	l Pro	Va1	Trp	Lys	s Glu	ı Ala	a Th:	Thi	Th	r Le	u Pho	e Cy
	Ala	a Sei	r Asp) Ala	Lys 20		Tyr	: Asp	Ser	Glu 25		His	s Ası	n Vai	l Tri
15	Ala	a Thi	: His	a Ala	Cys 35	Val	Pro	Thr	Asp	Pro 40		Pro	Glr	Glu	Va! 45
	Glu	ı Leu	; Glu	Asn	Val 50	Thr	Glu	Asn	Phe	Asn 55		Trp	Lys	a Asr	Asr 60
20	Met	: Val	Glu	Gln	Met 65	His	Gly	Asp	Ile	70		Leu	Trp	Asp	Glr 75
25	Ser	Leu	Lys	Pro	Cys 80	Val	Lys	Leu	Thr	Pro 85		Cys	Val	Thr	Leu 90
	Asn	Cys	Thr	Asp	Pro 95	Asn	Val	Thr	Asn	Ser 100	Glu	Arg	Thr	Ile	Glu 105
30					Lys 110					115					120
					Gln 125					130					135
35					Gly 140					145					150
40					Val 155					160					165
					Ile 170					175					180
45					Asp 185					190					195
					Val 200					205					210
50					Leu 215					220				-	225
55					Ala 230					235					240
					Glu 245					250					255
60					Lys 260					265					270
	Tyr	Ala	Thr	Gly	Asn 275	Ile	Ile	Gly	Asp	Ile 280	Arg	Gln	Ala		Cys 285

	Asn	Ile	Ser	Gly	Thr 290	Lys	Trp	Asn	Asp	Thr 295	Leu	Lys	Lys	Ile	Ala 300
5	Ile	Lys	Leu	Arg	Glu 305	Gln	Phe	Asn	Lys	Thr 310	Ile	Val	Phe	Asn	Gln 315
	Ser	Ser	Gly	Gly	Asp 320	Pro	Glu	Ile	Ala	Thr 325	Leu	Ser	Phe	Asn	330 Cys
10	Gly	Gly	Glu	Phe	Phe 335	Tyr	Cys	Asn	Ser	Thr 340	Gln	Leu	Phe	Asn	Ser 345
15	Thr	Trp	Asn	Ser	Thr 350	Gly	Ser	Asn	Asn	Thr 355	Lys	Gly	Asn	Asp	Thr 360
15	Ile	Thr	Leu	Pro	Cys 365	Arg	Ile	Arg	Gln	Ile 370	Ile	Asn	Met	Trp	Gln 375
20	Lys	Ile	Gly	Lys	Ala 380	Met	Tyr	Ala	Pro	Pro 385	Ile	Lys	Gly	Gln	11e 390
	Arg	Cys	Ser	Ser	Asn 395	Ile	Thr	Gly	Leu	11e 400	Leu	Thr	Arg	Asp	Gly 405
25	Gly	Asn	Asn	Asn	Met 410	Ser	Lys	Thr	Thr	Glu 415	Thr	Phe	Arg	Pro	Gly 420
	Gly	Gly	Asp	Met	Arg 425	Asp	Asn	Trp	Arg	Ser 430	Glu	Leu	Tyr	Lys	Tyr 435
30	Lys	Val	Val	Lys	11e 440	Glu	Pro	Leu	Gly	Val 445	Ala	Pro	Thr	Arg	Ala 450
35	Lys	Arg	Arg	Val	Val 455	Gln	Arg	Glu	Lys	Arg 460	Ala	Val	Gly	Ile	Gly 465
	Ala	Val	Phe	Leu	Gly 470	Phe	Leu	Gly	Ala	Xaa 475	Ser	Phe	Xaa 478		
40	(2)] (i	L) SI () ()	EQUE! A) LI B) T	NCE (ENGT! (PE:	FOR SCHARA H: 14 Nucl	ACTE 135 l Leic	RIST: pase Acid	ICS: pai:							
45	(x:	(I i) SI	D) TO	NCE (OGY: DESCI	Line RIPT	ear ION:	SEQ							
50		CTC Leu 1	GAG Glu	GTA Val	CCT Pro	GTG Val 5	Trp	Lys	Glu	GCA Ala	Thr	ACC Thr	ACT Thr	36	
	CTA Leu	TTT Phe	TGT Cys 15	GCA Ala	TCA Ser	GAT Asp	GCT Ala	AAA Lys 20	GCA Ala	TAT Tyr	GAT Asp	TCA Ser	GAG Glu 25	75	
55	GCA Ala	CAT His	AAT Asn	Val	TGG Trp 30	GCC Ala	ACA Thr	CAT His	GCC Ala	TGT Cys 35	GTA Val	ccc Pro	ACA Thr	114	
60	GAC Asp	CCC Pro 40	Asn	CCA Pro	CAA Gln	GAA Glu	GTA Val 45	GAA Glu	TTG Leu	GAA Glu	AAT Asn	GTG Val 50	ACA Thr	153	

	GA/ Glu	AA A	T TT	T AAG Asi 5!	n Met	G TG	G AAI P Lys	A AA' 5 As:	T AAG n Asi 60	n Met	G GT	A GA l Gl	A CA u Gl	G 192 n
5	ATO Met 65	Hi	T GG(s Gl ₎	G GAT	T ATA	A AT:	e Sei	TTI Le	A TGC	G GAT	CAA Gli 75	n Se	C CT.	A 231
10	AAC Lys	CC	A TGT Cys 80	3 Val	A AAA L Lys	TT/	A ACC	CCP Pro) Lev	TGT Cys	GT1	T AC	TTI r Leu 90	
15	AA1 Asn	TG(C ACT	GAC Asp	CCA Pro 95) Asr	GTI Val	ACT Thr	TAA ? Asn	Ser 100	Glu	AG/	A ACC	309
20	ATA Ile	GAC Glu 105	ı Gly	GGA Gly	GAA Glu	ATA Ile	AAA Lys 110	Asn	TGC Cys	TCT Ser	TTC Phe	AA1 Asr 115	lle	348
	ACC Thr	ACA Thr	AAC Asn	Ile 120	Arg	GAT Asp	AGG Arg	TTT Phe	CAG Gln 125	AAA Lys	GAA Glu	TAT	GCA Ala	387
25	CTT Leu 130	Phe	TAT Tyr	AAA Lys	CTT Leu	GAT Asp 135	Val	ATA Ile	CCA Pro	TTA Leu	GGT Gly 140	AAT Asn	GAT Asp	426
30	AAT Asn	ACT Thr	AGC Ser 145	Tyr	AGG Arg	TTG Leu	ATA Ile	AGT Ser 150	TGT Cys	AAC Asn	ACC Thr	TCA Ser	GTC Val 155	465
35	ATT Ile	ACA Thr	CAG Gln	GCC Ala	TGT Cys 160	CCA Pro	AAG Lys	GTA Val	TCC Ser	TTT Phe 165	GAG Glu	CCA Pro	ATT Ile	504
40	CCC Pro	ATA Ile 170	CAT His	TAT Tyr	TGT Cys	GCC Ala	CCG Pro 175	GCT Ala	GGT Gly	TTT Phe	GCG Ala	ATT Ile 180	CTA Leu	543
	AAG Lys	TGT Cys	AAA Lys	GAT Asp 185	AAG Lys	AAG Lys	TTC Phe	AAT Asn	GGA Gly 190	ΛCA Thr	GGA Gly	CCA Pro	TGT Cys	582
45	ACA Thr 195	AAT Asn	GTC Val	AGC Ser	ACA Thr	GTA Val 200	CAA Gln	TGT Cys	ACA Thr	CAT His	GGA Gly 205	ATT	AAG Lys	621
50	CCA Pro	GTA Val	GTA Val 210	TCA Ser	ACT Thr	CAA Gln	CTG Leu	TTG Leu 215	TTA Leu	AAT Asn	GGC Gly	AGT Ser	CTA Leu 220	660
55	GCA Ala	GAA Glu	GAA Glu	GAC Asp	ATA Ile 225	GTA Val	ATT Ile	AGA Arg	TCC Ser	GCC Ala 230	AAT Asn	CTC Leu	ACA Thr	699
60	GAC Asp	AAT Asn 235	GCT Ala	AAA Lys	AAC Asn	ATA Ile	ATA Ile 240	GTA Val	CAG Gln	CTG Leu	AAT Asn	GAA Glu 245	TCT Ser	738
	GTA Val	ACA Thr	ATG Met	AAT Asn 250	TGT Cys	ACA Thr	AGA Arg	Pro	AAC . Asn . 255	AAC Asn	AAT Asn	ACA Thr	ATG Met	777

	AAA Lys 260	AGT Ser	ATA Ile	CAT His	ATA Ile	GGA Gly 265	CCA Pro	GGC Gly	AGA Arg	GCA Ala	TTT Phe 270	TAT Tyr	GCA Ala	816
5	ACA Thr	GGA Gly	AAC Asn 275	ATA Ile	ATA Ile	GGA Gly	GAT Asp	ATA Ile 280	AGA Arg	CAA Gln	GCA Ala	CAT His	TGT Cys 285	855
10	AAC Asn	ATT Ile	AGT Ser	GGA Gly	ACA Thr 290	AAA Lys	TGG Trp	AAT Asn	GAC Asp	ACT Thr 295	TTG Leu	AAA Lys	AAG Lys	894
15	ATA Ile	GCT Ala 300	ATA Ile	AAA Lys	TTA Leu	AGA Arg	GAA Glu 305	CAA Gln	TTT Phe	AAT Asn	AAG Lys	ACA Thr 310	ATA Ile	933
	GTC Val	TTT Phe	AAT Asn	CAA Gln 315	TCC Ser	TCA Ser	GGA Gly	GGG Gly	GAC Asp 320	CCA Pro	GAA Glu	ATT Ile	GCA Ala	972
20	ACG Thr 325	CTC Leu	AGT Ser	TTT Phe	AAT Asn	TGT Cys 330	GGA Gly	GGG Gly	GAA Glu	TTT Phe	TTC Phe 335	TAC Tyr	TGT Cys	1011
25	AAT Asn	TCA Ser	ACA Thr 340	CAA Gln	CTG Leu	TTT Phe	AAT Asn	AGT Ser 345	ACT Thr	TGG Trp	AAT Asn	AGT Ser	ACT Thr 350	1050
30	GGG Gly	TCA Ser	AAT Asn	AAC Asn	ACT Thr 355	AAA Lys	GGA Gly	AAT Asn	GAC Asp	ACA Thr 360	ATC Ile	ACA Thr	CTC Leu	1089
35	CCA Pro	TGC Cys 365	AGA Arg	ATA Ile	AGA Arg	CAA Gln	ATT Ile 370	ATA Ile	AAC Asn	ATG Met	TGG Trp	CAG Gln 375	AAA Lys	1128
	ATA Ile	GGA Gly	AAA Lys	GCA Ala 380	ATG Met	TAT Tyr	GCC Ala	CCT Pro	CCC Pro 385	ATC Ile	AAA Lys	GGG Gly	CAA Gln	1167
40	ATT Ile 390	AGA Arg	TGT Cys	TCA Ser	TCA Ser	AAT Asn 395	ATT Ile	ACA Thr	GGG Gly	CTA Leu	ATA Ile 400	TTA Leu	ACA Thr	1206
45	AGA Arg	GAT Asp	GGT Gly 405	GGT Gly	AAC Asn	AAC Asn	AAC Asn	ATG Met 410	AGC Ser	AAG Lys	ACC Thr	ACC Thr	GAG Glu 415	1245
50	ACC Thr	TTC Phe	AGA Arg	CCT Pro	GGA Gly 420	GGA Gly	GGA Gly	GAT Asp	ATG Met	AGG Arg 425	GAC Asp	AAT Asn	TGG Trp	1284
55	AGA Arg	AGT Ser 430	GAA Glu	TTA Leu	TAT Tyr	Lys	TAT Tyr 435	AAA Lys	GTA Val	GTA Val	AAA Lys	ATT Ile 440	GAA Glu	1323
	CCA Pro	TTA Leu	GGA Gly	GTA Val 445	GCA Ala	CCC Pro	ACC Thr	AGG Arg	GCA Ala 450	AAG Lys	AGA Arg	AGA Arg	GTG Val	1362
60	GTG Val 455	CAG Gln	AGA Arg	GAA Glu	AAA Lys	AGA Arg 460	GCA Ala	GTG Val	GGA Gly	ATA Ile	GGA Gly 465	GCT Ala	GTG Val	1401

TTC CTT GGG TTC TTG GGA GCA TAA AGC TTC TAG A 1435 Phe Leu Gly Phe Leu Gly Ala Xaa Ser Phe Xaa 475 5 (2) INFORMATION FOR SEQ ID NO:28: (i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 478 amino acids (B) TYPE: Amino Acid (D) TOPOLOGY: Linear 10 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:28: Leu Glu Val Pro Val Trp Lys Glu Ala Thr Thr Thr Leu Phe Cys 15 Ala Ser Asp Ala Lys Ala Tyr Asp Ser Glu Ala His Asn Val Trp Ala Thr His Ala Cys Val Pro Thr Asp Pro Asn Pro Gln Glu Val 20 Glu Leu Glu Asn Val Thr Glu Asn Phe Asn Met Trp Lys Asn Asn Met Val Glu Gln Met His Gly Asp Ile Ile Ser Leu Trp Asp Gln 25 Ser Leu Lys Pro Cys Val Lys Leu Thr Pro Leu Cys Val Thr Leu 30 Asn Cys Thr Asp Pro Asn Val Thr Asn Ser Glu Arg Thr Ile Glu 100 Gly Gly Glu Ile Lys Asn Cys Ser Phe Asn Ile Thr Thr Asn Ile 35 Arg Asp Arg Phe Gln Lys Glu Tyr Ala Leu Phe Tyr Lys Leu Asp 125 Val Ile Pro Leu Gly Asn Asp Asn Thr Ser Tyr Arg Leu Ile Ser 40 Cys Asn Thr Ser Val Ile Thr Gln Ala Cys Pro Lys Val Ser Phe 45 Glu Pro Ile Pro Ile His Tyr Cys Ala Pro Ala Gly Phe Ala Ile Leu Lys Cys Lys Asp Lys Lys Phe Asn Gly Thr Gly Pro Cys Thr 185 50 Asn Val Ser Thr Val Gln Cys Thr His Gly Ile Lys Pro Val Val 200 205 Ser Thr Gln Leu Leu Asn Gly Ser Leu Ala Glu Glu Asp Ile 55 Val Ile Arg Ser Ala Asn Leu Thr Asp Asn Ala Lys Asn Ile Ile 230 60 Val Gln Leu Asn Glu Ser Val Thr Met Asn Cys Thr Arg Pro Asn

265

Asn Asn Thr Met Lys Ser Ile His Ile Gly Pro Gly Arg Ala Phe

65

```
Tyr Ala Thr Gly Asn Ile Ile Gly Asp Ile Arg Gln Ala His Cys
       Asn Ile Ser Gly Thr Lys Trp Asn Asp Thr Leu Lys Lys Ile Ala
5
       Ile Lys Leu Arg Glu Gln Phe Asn Lys Thr Ile Val Phe Asn Gln
                       305
                                           310
       Ser Ser Gly Gly Asp Pro Glu Ile Ala Thr Leu Ser Phe Asn Cys
10
                                                                330
       Gly Gly Glu Phe Phe Tyr Cys Asn Ser Thr Gln Leu Phe Asn Ser
                                                                345
15
       Thr Trp Asn Ser Thr Gly Ser Asn Asn Thr Lys Gly Asn Asp Thr
                                                                360
       Ile Thr Leu Pro Cys Arg Ile Arg Gln Ile Ile Asn Met Trp Gln
20
       Lys Ile Gly Lys Ala Met Tyr Ala Pro Pro Ile Lys Gly Gln Ile
       Arg Cys Ser Ser Asn Ile Thr Gly Leu Ile Leu Thr Arg Asp Gly
25
       Gly Asn Asn Asn Met Ser Lys Thr Thr Glu Thr Phe Arg Pro Gly
                                           415
30
       Gly Gly Asp Met Arg Asp Asn Trp Arg Ser Glu Leu Tyr Lys Tyr
                                            430
       Lys Val Val Lys Ile Glu Pro Leu Gly Val Ala Pro Thr Arg Ala
                                            445
35
       Lys Arg Arg Val Val Gln Arg Glu Lys Arg Ala Val Gly Ile Gly
       Ala Val Phe Leu Gly Phe Leu Gly Ala Xaa Ser Phe Xaa
40
                                           475
                       470
      (2) INFORMATION FOR SEQ ID NO:29:
         (i) SEQUENCE CHARACTERISTICS:
             (A) LENGTH: 511 amino acids
45
             (B) TYPE: Amino Acid
(D) TOPOLOGY: Linear
         (xi) SEQUENCE DESCRIPTION: SEQ ID NO:29:
       Met Arg Val Lys Gly Ile Arg Arg Asn Tyr Gln His Trp Trp Gly Arg
50
       Gly Thr Met Leu Leu Gly Leu Leu Met Ile Cys Ser Ala Thr Glu Lys
55
       Leu Trp Val Thr Val Tyr Tyr Gly Val Pro Val Trp Lys Glu Ala Thr
       Thr Thr Leu Phe Cys Ala Ser Asp Ala Lys Ala Tyr Asp Thr Glu Ala
60
       His Asn Val Trp Ala Thr His Ala Cys Val Pro Thr Asp Pro Asn Pro
```

	G1	n Gl	lu Va	l Gl	u Lei 85	u Va	l As	n Va	1 Th	r G1	lu As O	sn P	he A	sn	Met	Trp 95	Lys
5	As	n As	n Me	t Va 10	l Glu O	ı Gl	n Mei	t Hi	s G1 10	u As	sp I	le I	le S	er	Leu 110	Trp	Asn
	Gl	n Se	r Le 11	u Ly: 5	s Pro	су:	s Va	1 Ly:	s Le O	u Th	r Pr	o Le		ys 25	Val	Thr	Leu
10	As	n Cy 13	s Th O	r Ası	p Leu	Ar	g Asr 139	n Thi	r Th	r As	n Th	r As	n A	sn .	Ser	Thr	Asp
15	As:	n As 5	n As	n Sez	c Lys	Se:	r Glu	ı Gly	/ Th	r Il	e Ly 15	s G1 5	y G	ly (Glu	Met	Lys 160
	Ası	n Cy	s Se	r Phe	2 Asn 165	Ile	? Thr	Thr	: Se	r Il 17	e G1 O	y As	рL	ys l	Met	Gln 175	Lys
20	Gla	ту:	r Ala	180	Leu	Tyr	Lys	Leu	189	o Il	e Gl	u Pr	0 I		Asp 190	Asn	Asp
	Se	Th	r Ser 195	Tyr	Arg	Leu	Ile	Ser 200	Cys	A A S	n Th	r Se	r Va 20		le	Thr	Gln
25	Ala	210	Pro	Lys	Ile	Ser	Phe 215	Glu	Pro	Ile	Pro	22	e Hi O	s T	'yr	Cys	Ala
30	227	*			Ala	230					235	,					240
	Lys	Gly	/ Ser	Cys	Lys 245	Asn	Val	Ser	Thr	Val 250	Glr	n Cys	Th	r H		31y 255	Ile
35			280		Ser				265					2	70		
	Glu	Glu	Val 275	Val	Ile	Arg	Ser	Glu 280	Asp	Phe	Thr	Asp	As 28		la I	ys '	Thr
40	Ile	Ile 290	Val	His	Leu	Lys	Glu 295	Ser	Val	Gln	Ile	Asn 300	Су	s Ti	nr A	irg 1	Pro
45	Asn 305	Tyr	Asn	Lys	Arg	Lys 310	Arg	Ile	His	Ile	Gly 315	Pro	Gl	y Az	cg A		Phe 320
					Asn 325					330					3	35	
50				340	Lys				345					35	0		_
			333		Phe :			360					365	1			
55	Gly	Gly 370	Asp	Pro	Glu :	lle	Val 375	Met	His	Ser	Phe	Asn 380	Cys	Gl	y G	ly G	lu
60	303				-	390					395					4	00
					Asn # 405					410					4]	۱5	
65	Gln	Cys	Lys	Ile 1 420	Lys (Sln .	Ile :	Ile i	Asn	Met	Trp	Gln	Lys	Va.	1 G1	y L	ys

	Ala Met	Tyr 7	Ala Pr	o Pro	11	Glu 440	Gly	Gln	Ile	Arg	Cys 445	Ser	Ser	Asn
5	Ile Thr 450	Gly I	Leu Le	u Leu	Thr 455	Arg	Asp	Gly	Gly	Glu 460	Asp	Thr	Asp	Thr
	Asn Asp 465	Thr C	3lu Il	e Phe 470	Arg	Pro	Gly	Gly	Gly 475	Asp	Met	Arg	Asp	Asn 480
10	Trp Arg	Ser 0	Glu Le 48		Lys	Tyr	Lys	Val 490	Val	Thr	Ile	Glu	Pro 495	Leu
15	Gly Val		Pro Th	r Lys	Ala	Lys	Arg 505	Arg	Val	Val	Gln	Arg 510	Glu	
20		(A) (B) (C)	ON FOR ENCE C LENGT TYPE: STRAN TOPOL ENCE D	HARACT H: 280 Nucle DEDNES OGY: I	renis 00 ba eic A ss: s Linea	STICS ase p Acid Singl	oairs Le		30:					
25	TTCGAGCT	cc ccc	CGACAT	TG ATT	OTTAT	SACT	AGAC	TCGA	ATC C	CACAC	CTG	rG	50	
	GAATGTGT	GT CAC	STTAGG	GT GT	GAAA	AGTC	CCCA	GGCI	rcc c	CAGO	AGG	CA	100)
• •	GAAGTATG	CA AAC	CATGC	AT CTO	CAATI	TAGT	CAGO	CAACC	CAG C	TGTO	GAA	AG	150	t
30	TCCCCAGG	CT CCC	CAGCA	GG CAC	GAAGT	TATG	CAAA	GCAT	GC F	TCTC	CTAAT	CA.	20	0
	GTCAGCAA	CC AT	AGTCCC	GC CCC	CTAAC	CTCC	GCCC	CATCO	CCG	cccc	TAAC	CTC	25	0
35	CGCCCAGT	TC CGC	CCATT	ст ссс	cccc	CATG	GCTG	ACTA	AT 1	TTTT	TTAT	T	30	0
	TATGCAGA	GG CCC	GAGGCC	GC CTC	GGCC	CTCT	GAGC	TATI	CC P	GAAC	TAGI	rG	35	0
	AGGAGGCT"	TT TT	rggagg	CC TAC	GCTI	TTG	CAAA	AAGC	TA C	CTT	TCCC	G	400	
40	CCGGGAAC	GG TGC	CATTGG	AA CG	CGGAT	TCC	CCGT	CCA	AG A	GTC	AGGT#	A.	450	
	GTACCGCC	TA TAC	GAGTCT	AT AGO	CCCF	ACCC	CCTI	GGCI	TC C	TTAC	AACC	C	500	•
45	GGCTACAA	TT AAT	racata	AC CT	rttgo	ATC	GATO	CTAC	TG A	CAC1	GAC	ΑT	550	
	CCACTTTT	TC TT	TTTCTC	CA CAC	GTGI	CCA	СТСС	CAGO	TC (CAACI	GCAC	cc	600	ŀ
	TCGGTTCG	CG AAG	GCTAGC	TT GG	CTG	CATC	GATI	GAAT	TC C	CACTO	CCTI	C	650	
50	CACCAAGC	TC TG	CAGGAT	CC CAC	GAGTO	CAGG	GG 1	CT C	TA 7	CT Ter S	cc 1 er C	rgc Cys	69	7
								1				5		•
55	TGG TGG Trp Trp	CTC C Leu G	AG TTC ln Phe 10	AGG A	AAC A Asn S	AGT <i>I</i> Ser I	AAA C Lys F	cc 1 ro 0	rgc 1	cc 6	SAA 1 Slu 1	TAT Tyr	73	9
60	TGC CTC Cys Leu 20	TCA CA Ser H:	AT CTC is Leu	GTC I Val I 25	AAT (Asn I	CTC (Leu /	cgc c	AG C	SAC 1 Asp 1	rgG C	GA C	ccc	78	1

	TCT (Cys)	GAC A Asp I 35	AG CI ys Le	T CAC	G CGC	GAA Glu 40	Arc	CCA Pro	A ACT	r Acc	c cc r Pr 4	o Il	C ATC e Ile	823
5	AGT 1	rat c ryr P	CT TA ro * 50	A GGT	CTC Leu	TTT Phe	TG1 Cys 55	Val	GTC Val	G CG1	r TC	C GG r G1 6	T ATG y Met O	865
10	62 62	Hy T	hr Al 6	a Ala 5	Arg	, Leu	Gly	Ala 70	Val	Ile	e Le	ı Phe	75	907
15	GTC A	TA G	TG GG al Gl	C CTC y Leu 80	His	GGG	GTC Val	CGC Arg	GGC Gly 8	Lys	TAT	GCC Ala	C TTG	949
20	GCG G Ala A 90	AT G	CC TC la Se	r CTC r Leu	AAG Lys 95	ATG Met	GCC Ala	GAC Asp	CCC Pro	AAT Asn 100	Arg	TTI Phe	CGC Arg	991
25	GGC A Gly L	ys As	sp Lei)5	ı Pro	Val	Leu	Asp 110	Gln	Leu	Leu	Glu	Val 115	Pro	1033
	GTG TO	GG AA	A GAA	A GCA Ala 120	AAC Asn	ACC Thr	ACT Thr 125	CTA Leu	TTT Phe	TGT Cys	GCA Ala	TCA Ser 130	Asp	1075
30	GCT A	AA GO ys Al	A TAT a Tyr 135	. Lys	ACA Thr	GAG Glu	GCA Ala	CAT His 140	AAT Asn	GTT Val	TGG Trp	GCC Ala	ACA Thr 145	1117
35	CAT GO His A	la Cy	s Val	150	Thr	Asp	Pro	Lys	Pro 155	Gln	Glu	Ile	Lys	1159
40.	TTG GA Leu Gl 160	lu As	n Val	Thr	Glu 165	Asn	Phe	Asn	Met	Trp 170	Lys	Asn	Asn	1201
45	ATG GT Met Va 17	11 G1 75	u Gln	Met	His	Glu 180	Asp	Ile	Ile	Ser	Leu 185	Trp	Asp	1243
	CAA AG Gln Se	r Le 19	u Lys D	Pro	Cys	Val :	Lys 195	Leu	Thr	Pro	Leu	Cys 200	Val	1285
50	ACT TT Thr Le	'A AA' 'u As	T TGC Cys 205	ACT Thr	Aab GYL	TTG /	Arg A	AAT A Asn a 210	AAT . Asn '	ACT Thr	AAT Asn	ACC Thr	AAT Asn 215	1327
5 5	AGT AC Ser Th	TAC TY:	GGA Gly	AAA Lys 220	ATA I	ATG (Met (GAG (Glu (3ly (GGA (Gly (225	GAG . Glu	ATA Ile	AAA Lys	AAC Asn	1369
60	TGC TC Cys Se 230	T TTO r Phe	AAT Asn	Ile '	ACC I Thr 1 235	ACA A	AGC A Ser 1	NTA A	Lys A	SAT A Asp 1	AAG Lys	CTG Leu	AAA Lys	1411
65	GAT ATO Asp Me 24	t Se	CTT Leu	TTT '	fyr I	AAA C Lys I 250	eu A	AT C	GTA C	al E	CCA Pro 255	ATA (GGT Gly	1453

_	TAA Asn	AAT Asn	AGT Ser 260	AAT Asn	ACT Thr	ACT Thr	AGT Ser	TAT Tyr 265	AGG Arg	TTG Leu	ATA Ile	AGT Ser	TGT Cys 270	AAC Asn	1495
5														GAG Glu 285	1537
10	CCA Pro	ATT Ile	CCC Pro	ATA Ile	CAT His 290	TAT Tyr	TGT Cys	GCC Ala	CCG Pro	GCT Ala 295	GGT Gly	TTT Phe	GCG Ala	ATT Ile	1579
15	CTC Leu 300	AAG Lys	TGT Cys	AAT Asn	GAT Asp	AAT Asn 305	AAG Lys	TTC Phe	AAT Asn	GGA Gly	ACA Thr 310	GGA Gly	CCA Pro	TGT Cys	1621
20	CCA Pro	AAT Asn 315	GTC Val	AGC Ser	ACA Thr	GTA Val	CAA Gln 320	TGT Cys	ACA Thr	CAT His	GGA Gly	ATT Ile 325	AGG Arg	CCA Pro	1663
25	GTA Val	GTA Val	TCA Ser 330	ACT Thr	CAA Gln	CTG Leu	CTG Leu	TTA Leu 335	AAT Asn	GGC Gly	AGT Ser	CTA Leu	GCA Ala 340	GAA Glu	1705
25	AAA Lys	GAG Glu	GTA Val	GTC Val 345	CTT Leu	AGA Arg	TCT Ser	GAA Glu	AAT Asn 350	TTC Phe	ACG Thr	GAC Asp	AAT Asn	GCT Ala 355	1747
30	AAA Lys	ACC Thr	ATA Ile	ATA Ile	GTA Val 360	CAG Gln	CTG Leu	AAC Asn	GAA Glu	TCT Ser 365	GTA Val	ATA Ile	ATT	GAT Asp	1789
35	TGT Cys 370	ATG Met	AGA Arg	CCC Pro	AAC Asn	AAC Asn 375	AAT Asn	ACA Thr	AGA Arg	ACA Thr	AGT Ser 380	ATA Ile	CCT Pro	ATG Met	1831
40	GGA Gly	CCA Pro 385	GGG Gly	AAA Lys	GCA Ala	TTT Phe	TAT Tyr 390	GCA Ala	ACA Thr	GGA Gly	GAT Asp	GTA Val 395	ATA Ile	GGA Gly	1873
45	GAT Asp	ATA Ile	AGA Arg 400	CGA Arg	GCA Ala	CAT His	CAa LCL	AAC Asn 405	ATT Ile	AGT Ser	AGA Arg	GCA Ala	GGA Gly 410	TGG Trp	1915
45	AAT Asn	ACC Thr	ACT Thr	TTA Leu 415	CAA Gln	CAG Gln	ATA Ile	GCT Ala	AAA Lys 420	AAA Lys	TTA Leu	AGA Arg	GAA Glu	AAA Lys 425	1957
50											TCC Ser				1999
55	GAC Asp 440	CCA Pro	GAA Glu	ATT Ile	GTA Val	ATG Met 445	CAC His	ACT Thr	TTT Phe	AAT Asn	TGT Cys 450	GGA Gly	GGG Gly	GAA Glu	2041
60	TTT Phe	TTC Phe 455	TGC Cys	TGT Cys	AAT Asn	TCA Ser	ACA Thr 460	CCA Pro	CTG Leu	TTT Phe	AAT Asn	AGT Ser 465	ACT Thr	TGG Trp	2083
65	AAT Asn	GAT Asp	GCA Ala 470	CAA Gln	CTG Leu	TTT Phe	AAT Asn	AGT Ser 475	ACT Thr	TGG Trp	GAT Asp	GAT Asp	ACT Thr 480	AAA Lys	2125

	TGG	TCA Ser	AAA Lys	GGC Gly 485	ACT Thr	AAC Asn	GAA Glu	AAT Asn	GAC Asp 490	Thr	ATC	ACC Thr	CTC Leu	CAT His 495	2167
5	TGC Cys	AGA Arg	ATA Ile	AAA Lys	CAA Gln 500	Ile	ATA Ile	AAT Asn	ATG Met	TGG Trp 50	Gln	GAA Glu	GTA Val	GGA Gly	2209
10	AAA Lys 510	Ala	ATG Met	TAT Tyr	GCC Ala	CCT Pro 515	Pro	ATC Ile	AAA Lys	GGA Gly	CAA Gln 520	ATT Ile	AGA Arg	TGT Cys	2251
	GAA	TCA	AAT	ATT	ACA	GGG	CTG	CTA	TTA	ACA	AGA	GAT	GGT	GGT	2293
15	Glu	Ser 525	Asn	Ile	Thr	Gly	Leu 530	Leu	Leu	Thr	Arg	Asp 535	Gly	Gly	
20	AAC Asn	GAC Asp	ACG Thr 540	AGC Ser	AAG Lys	AAT Asn	AAC Asn	ACT Thr 545	GAG Glu	ATT Ile	TTC Phe	AGA Arg	CCT Pro 550	GGA Gly	2335
25	GGA Gly	GGA Gly	AAT Asn	ATG Met 555	AAG Lys	GAC Asp	AAT Asn	TGG Trp	AGA Arg 560	AGT Ser	GAA Glu	TTA Leu	TAT Tyr	AAA Lys 565	2377
30	TAT Tyr	AAA Lys	GTA Val	Ile	AAA Lys 570	ATT Ile	GAA Glu	CCA Pro	TTA Leu	GGA Gly 575	GTA Val	GCA Ala	CCC Pro	ATC Ile 579	2419
	TAGG	CAAA	GA G	AAGA	GTGG	T GC	AGAG	AGAA	AAA	AGAG	CAG	TGAC	ACTA	GG	2469
35	AGCT	'ATGT	TC C	TTGG	GTTC	T TG	GGAG	CAGC	AGG	AAGC	лст	ATGG	GCGA	TA	2519
	AGCT	'TTAA	TG C	GGTA	GTTT	A TC	ACAG	TTAA	ATT	CGTA	ACG (CAGT	CAGG	CA	2569
40	CCGT	GTAT	GA A	ATCT.	AACA	A TG	CGAC	CTGC	AGA	AGCT'	TAG A	AACC	GAGG.	AA	2619
	CTTG	TTTA	TT G	CAGC	TTATA	ATA F	GGTT	ACAA	ATA.	AAGC	AAT I	AGCA:	rcac.	AA	2669
	ATTT	CACA	AA T	AAAG	CATTI	TT T	TTCA	CTGC	ATT	CTAG	TTG :	rggt:	TTGT	CC	2719
45	AAAC	TCAT	CA A	TGTA:	rctt <i>i</i>	TC	ATGT	CTGG	ATC	GGGA	TT A	TTAF	cggc	SC	2769
	AGCA	CCAT	GG C	CTGA	ATA	CC	TCTG	AAAG	A						2800
50	(2)	(i) SI	EQUEN (A) (B) (D)	FOR SINCE OF LENGTYPE	HARI TH: : An	ACTER 579 mino Y: Li	RISTI amin Acio inear	no ac I						
55					ICE D										
	Ser v	val :	ser S	ser C	ys T 5	rp 1	rp L	.eu G	iln F	he A	rg A	sn S	er L		ro Cys 15
60	Ser (Glu :	ryr c	ys L 20	eu S	er H	dis L	.eu V	al A 25	sn L	eu A	rg G	lu A	sp T:	rp Gly
65	Pro (Cys #	Asp L 35	ys L	eu G	ln A	Arg G	lu A 40	rg P	ro T	hr T		ro I 45	le I	le Ser

	Tyr	Pro 50	*	Gly	Leu	Phe	Cys 55	Val	Val	Arg	Ser	Gly 60		Gly	Gly	Thr
5	Ala 65	Ala	Arg	Leu	Gly	Ala 70	Val	Ile	Leu	Phe	Val 75	Val	Ile	Val	Gly	Leu 80
	His	Gly	Val	Arg	Gly 85	Lys	Tyr	Ala	Leu	Ala 90	Asp	Ala	Ser	Leu	Lys 95	Met
10	Ala	Asp	Pro	Asn 100	Arg	Phe	Arg	Gly	Lys 105	Asp	Leu	Pro	Val	Leu 110	Aap	Gln
15	Leu	Leu	Glu 115	Val	Pro	Val	Trp	Lys 120	Glu	Ala	Asn	Thr	Thr 125	Leu	Phe	Cys
15	Ala	Ser 130	Asp	Ala	Lys	Ala	Tyr 135	Lys	Thr	Glu	Ala	His 140	Asn	Val	Trp	Ala
20	Thr 145	His	Ala	Cys	Val	Pro 150	Thr	Asp	Pro	Lys	Pro 155	Gln	Glu	Ile	Lys	Leu 160
	Glu	Asn	Val	Thr	Glu 165	Asn	Phe	Asn	Met	Trp 170	Lys	Asn	Asn	Met	Val 175	Glu
25	Gln	Met	His	Glu 180	Asp	Ile	Ile	Ser	Leu 185	Trp	Asp	Gln	Ser	Leu 190	Lys	Pro
30	Cys	Val	Lys 195	Leu	Thr	Pro	Leu	Cys 200	Val	Thr	Leu	Asn	Cys 205	Thr	Asp	Leu
50	Arg	Asn 210	Asn	Thr	Asn	Thr	Asn 215	Ser	Thr	Tyr	Gly	Lys 220	Ile	Met	Glu	Gly
35	225		Ile	-		230					235					240
	_		Lys		245					250					255	
40	_		Asn	260					265					270		
45			Ile 275					280					285			
		290	Tyr				295					300				
50	305		Phe			310					315					320
	_		His		325					330					335	
5 5	•	•	Leu	340					345					350		
60			Ala 355					360					365			
- -	_	370	Met				375					380	-			
65	Pro		Lys	Ala	Phe	Tyr 390	Ala	Thr	Gly	Asp	Val 395	Ile	Gly	Asp	Ile	Arg 400

	Arg	, Ala	His	Суз	Asn 405	Ile	Ser	Arg	Ala	Gly 410	Trp	Asn	Thr	Thr	Leu 415	Gln
5	Glr	Ile	Ala	Lys 420	Lys	Leu	Arg	Glu	Lys 425	Phe	Glu	Asn	Lys	Thr 430		Val
	Phe	. Asu	His 435	Ser	Ser	Gly	Gly	Asp 440	Pro	Glu	Ile	Val	Met 445	His	Thr	Phe
10	Asn	Cys 450	Gly	Gly	Glu	Phe	Phe 455	Cys	Cys	Asn	Ser	Thr 460	Pro	Leu	Phe	Asn
15	Ser 465	Thr	Trp	Asn	Asp	Ala 470	Gln	Leu	Phe	Asn	Ser 475	Thr	Trp	Asp	Asp	Thr 480
	Lys	Trp	Ser	Lys	Gly 485	Thr	Asn	Glu	Asn	Asp 490	Thr	Ile	Thr	Leu	His 495	Cys
20	Arg	Ile	Lys	Gln 500	Ile	Ile	Asn	Met	Trp 505	Gln	Glu	Val	Gly	Lys 510	Ala	Met
	Tyr	Ala	Pro 515	Pro	Ile	Lys	Gly	Gln 520	Ile	Arg	Cys	Glu	Ser 525	Asn	Ile	Thr
25	Gly	Leu 530	Leu	Leu	Thr	Arg	Asp 535	Gly	Gly	Asn		Thr 540	Ser	Lys	Asn	Asn
30	Thr 545	Glu	Ile	Phe	Arg	Pro 550	Gly	Gly	Gly	Asn	Met 555	Lys	Asp	Asn		Arg 560
	Ser	Glu	Leu	Tyr	Lys ' 565	Tyr	Lys	Val	Ile	Lys 570	Ile	Glu	Pro		Gly 575	Val
35	Ala	Pro	Ile													
40	(2)	<pre>INFO (i) (xi)</pre>	SEQI (A (B (C (D	UENCI) LEI) TYI) STI) TOI	E CHA NGTH: PE: A RANDE POLOC	ARAC' : 15 Nucle EDNE: SY: 1	TERI 33 b eic / SS: : Line	STIC: ase Acid Sing: ar	S: pair: le		:32:					
45	ATGG	GGGG	GA C	rgcco	CCAC	GT	rggg	GCC	GTG!	ATTT?	rgr 1	TGT	CGTC	AT.	50	
	AGTG	eecc.	TC C	ATGGC	GTCC	GCC	GCA	ATA	TGC	CTTGC	CG C	ATGO	стст	rc	100)
50		GATG												rg	150)
	GACC	AGCT	SC TO	GAG	GTA Val 1	CCT Pro	GTG Val	TGG Trp	AAA Lys 5	GAA Glu	GCA Ala	ACC Thr	ACC Thr		192	!
55	ACT Thr 10	CTA ? Leu I	TTT T	GT G	CA T la S	CA C er A 15	SAT C	CT A	AA G	CA T	AT G yr A 20	AT A	CA G	AG lu	23	4
60	GTA (CAT A His A 25	AT G	TT T	GG G rp A	CC A la T	CA C hr H	AT G	CC T la C	GT G ys V	al P	CC A ro T 35	CA G	AC sp	27	6

	CCC Pro	AAC Asn	CCA Pro 40	CAA Gln	GAA Glu	ATA Ile	GGA Gly	TTG Leu 45	GAA Glu	AAT Asn	GTA Val	ACA Thr	GAA Glu 50	AAT Asn	318
5	TTT Phe	AAC Asn	ATG Met	TGG Trp 55	FÀ2 YYY	AAT Asn	AAC Asn	ATG Met	GTA Val 60	GAA Glu	CAG Gln	ATG Met	CAT His	GAG Glu 65	360
10	GAT Asp	ATA Ile	ATC Ile	AGT Ser	TTA Leu 70	TGG Trp	GAT Asp	CAA Gln	AGC Ser	TTA Leu 75	AAG Lys	CCA Pro	TGT Cys	GTA Val	402
15	AAA Lys 80	TTA Leu	ACC Thr	CCA Pro	CTA Leu	TGT Cys 85	va:	ACT L Thi	TTA Lei	TAA IBA L	TGC Cy:	s Th	GAT Asj	TTG Leu	444
20	AAA Lys	AAT Asn 95	GCT Ala	ACT Thr	AAT Asn	ACC Thr	ACT Thr 100	AGT Ser	AGC Ser	AGC Ser	TGG Trp	GGA Gly 105	AAG Lys	ATG Met	486
	GAG Glu	AGA Arg	GGA Gly 110	GAA Glu	ATA Ile	AAA Lys	AAC Asn	TGC Cys 115	TCT Ser	TTC Phe	AAT Asn	GTC Val	ACC Thr 120	ACA Thr	528
25	AGT Ser	ATA Ile	AGA Arg	GAT Asp 125	AAG Lys	ATG Met	AAG Lys	AAT Asn	GAA Glu 130	TAT Tyr	GCA Ala	CTT Leu	TTT Phe	TAT Tyr 135	570
30	AAA Lys	CTT Leu	GAT Asp	GTA Val	GTA Val 140	CCA Pro	ATA Ile	GAT Asp	AAT Asn	GAT Asp 145	AAT Asn	ACT Thr	AGC Ser	TAT Tyr	612
35	AGG Arg 150	TTG Leu	ATA Ile	AGT Ser	TGT Cys	AAC Asn 155	ACC Thr	TCA Ser	GTC Val	ATT Ile	ACA Thr 160	CAG Gln	GCC Ala	TGT Cys	654
	CCA	AAG	GTG	TCC	TŢT	GAG	CCA	ATT	CCC	ATA	CAT	TAT	TGT	GCC	696
40		165			Phe		170					175			
45	CCG Pro	GCT Ala	GGT Gly 180	TTT Phe	GCG Ala	ATT	CTA Leu	AAG Lys 185	TGT Cys	AGA Arg	GAT Asp	AAA Lys	AAG Lys 190	TTC Phe	738
	AAC Asn	GGA Gly	ACA Thr	GGA Gly 195	CCA Pro	TGT Cys	ACA Thr	AAT Asn	GTC Val 200	AGC Ser	ACA Thr	GTA Val	CAA Gln	TGT Cys 205	780
50	ACA Thr	CAT His	GGA Gly	ATT Ile	AGG Arg 210	CCA Pro	GTA Val	GTA Val	TCA Ser	ACT Thr 215	CAA Gln	CTG Leu	CTG Leu	TTA Leu	822
55	AAT Asn 220	GGC Gly	AGT Ser	TTA Leu	GCA Ala	GAA Glu 225	GAA Glu	GAA Glu	GTA Val	GTA Val	ATT Ile 230	AGA Arg	TCT Ser	GCC Ala	864
60	AAT Asn	TTC Phe 235	TCG Ser	GAC Asp	AAT Asn	GCT Ala	AAA Lys 240	ACC Thr	ATA Ile	ATA Ile	GTA Val	CAG Gln 245	CTG Leu	AAC Asn	906

	GA: Gl:	A TC u Se	T GT r Va 25	l Gl	A AT	P AS	T TG n Cy	T AC s Th 25	r Ar	A CC g Pr	C AA O As	C AA n As	C A/ n As 26	AT ACA in Thr	948
5	AG/ Arc	A AG	A AG g Se	T ATA 116 265	∍ Hia	T ATA	A GG:	A CC. y Pro	A GG0 Gly 270	/ Ar	A GC g Al	A TT a Ph	T TA e Ty	T GCA r Ala 275	990
10	AC# Thr	GG;	A GAZ Y Glu	A ATA	A AT# 2 Ile 280	e Gly	A GAG / Asi	C ATA	A AGA	CA G1 28	n Al	A CA' a Hi	T TG s Cy	T AAC s Asn	1032
15	CTT Leu 290	Sei	C AGO	ACA Thr	AAA Lys	TGG Trp 295) Asr	CAA 1 naa r	ACT Thr	TT.	A AAA Lys	s Gli	G AT	A GTT e Val	1074
20	ACA Thr	Lys 305	Leu	AGA Arg	GAA Glu	CAT His	TTT Phe 310	Asn	'AAA Lys	AC/ Thi	ATA F	A GT0 ≥ Val 315	Ph	T AAT ≘ Asn	1116
25	CAC His	TCC	TCA Ser 320	Gly	GGG Gly	GAC Asp	CCA Pro	GAA Glu 325	Ile	GT# Val	λ ΛΤC . Met	CAC His	330	TTT Phe	1158
	AAT Asn	TGT Cys	GGA Gly	GGG Gly 335	GAA Glu	TTT Phe	TTC Phe	TAC	TGT Cys 340	AAT Asn	ACA Thr	ACA Thr	Pro	CTG Leu 345	1200
30	TTT Phe	AAT Asn	AGT Ser	ACT Thr	TGG Trp 350	AAT Asn	TAT Tyr	ACT Thr	TAT Tyr	ACT Thr 355	Trp	AAT Asn	raa naa	ACT Thr	1242
35	GAA Glu 360	GGG	TCA Ser	AAT Asn	GAC Asp	ACT Thr 365	GGA Gly	AGA Arg	AAT Asn	ATC Ile	ACA Thr 370	CTC Leu	CAA Gln	TGC Cys	1284
40	AGA Arg	ATA Ile 375	AAA Lys	CAA Gln	ATT Ile	ATA Ile	AAC Asn 380	ATG Met	TGG Trp	CAG Gln	GAA Glu	CTA Val 385	GGA Gly	AAA Lys	1326
	GCA	ATG	TAT	GCC	CCT	CCC	ATA	AGA	GGA	CAA	ATT	AGA	TGC	TCA	1368
45			Tyr 390					395					400		
50	TCA Ser	AAT Asn	ATT Ile	ACA Thr 405	GGG Gly	CTG Leu	CTA Leu	TTA Leu	ACA Thr 410	AGA Arg	GAT Asp	GGT Gly	GGT Gly	AAT Asn 415	1410
	AAC Asn	AGC Ser	GAA Glu	Thr	GAG Glu 420	ATC Ile	TTC Phe	AGA Arg	Pro	GGA Gly 425	GGA Gly	GGA Gly	GAT Asp	ATG Met	1452
5 5	AGG Arg 430	Aab Aab	AAT Asn	TGG Trp	Arg	AGT Ser 435	GAA Glu	TTA Leu	TAT Tyr	AAA Lys	TAT Tyr 440	AAA Lys	GTA Val	GTA Val	1494
60	AAA Lys	ATT Ile 445	GAA Glu	CCA Pro	TTA Leu	Gly	GTA Val 450	GCA Ala	CCC . Pro '	ACC Thr	Lys	GCA Ala 455	TAA *		1533

_	(2)	INF	ORMA:	SEQUI (A (B	ENCE) LEI) TYI	CHAI NGTH PE:	RACTI : 450 Amino	ERIS' 6 am 0 Ac	rics ino a id	: acid:	s					
5		(:	ki) :) TOI ENCE					Q ID	NO:	33:				
	Val 1	Pro	Val	Trp	Lys 5	Glu	Ala	Thr	Thr	Thr 10	Leu	Phe	Cys	Ala	Ser 15	Asp
10	Ala	Lys	Ala	Tyr 20	Asp	Thr	Glu	Val	His 25	Asn	Val	Trp	Ala	Thr 30	His	Ala
15	САа	Val	Pro 35	Thr	Asp	Pro	Asn	Pro 40	Gln	Glu	Ile	Gly	Leu 45	Glu	Asn	Val
	Thr	Glu 50	Asn	Phe	Asn	Met	Trp 55	Lys	Asn	Asn	Met	Val 60	Glu	Gln	Met	His
20	Glu 65	Asp	Ile	Ile	Ser	Leu 70	Trp	Asp	Gln	Ser	Leu 75	Lys	Pro	Cys	Val	Lys 80
25	Leu	Thr	Pro	Leu	Cys 85	Val	Thr	Leu	Asn	Cys 90	Thr	Asp	Leu	Lys	Asn 95	Ala
25	Thr	Asn	Thr	Thr 100	Ser	Ser	Ser	Trp	Gly 105	Lys	Met	Glu	Arg	Gly 110	Glu	Ile
30	Lys	Asn	Cys 115	Ser	Phe	Asn	Val	Thr 120	Thr	Ser	Ile	Arg	Asp 125	Lys	Met	Lys
	Asn	Glu 130	Туr	Ala	Leu	Phe	Tyr 135	Lys	Leu	Asp	Val	Val 140	Pro	Ile	Asp	Asn
35	Asp 145	Asn	Thr	Ser	Tyr	Arg 150	Leu	Ile	Ser	Cys	Asn 155	Thr	Ser	Val	Ile	Thr 160
40	Gln	Ala	Cys	Pro	Lys 165	Val	Ser	Phe	Glu	Pro 170	Ile	Pro	Ile	His	Tyr 175	Cys
40	Ala	Pro	Ala	Gly 180	Phe	Ala	Ile	Leu	Lys 185	Cys	Arg	Asp	Lys	Lys 190	Phe	Asn
45	Gly	Thr	Gly 195	Pro	CÀa	Thr	Asn	Val 200	Ser	Thr	Val	Gln	Cys 205	Thr	His	Gly
	Ile	Arg 210	Pro	Val	Val	Ser	Thr 215	Gln	Leu	Leu	Leu	Asn 220	Gly	Ser	Leu	Ala
50	Glu 225	Glu	Glu	Val	Val	1le 230		Ser	Ala	Asn	Phe 235	Ser	Asp	Asn	Ala	Lys 240
EE	Thr	Ile	Ile	Val	Gln 245	Leu	Asn	Glu	Ser	Val 250	Glu	Ile	Asn	Cys	Thr 255	Arg
55	Pro	Asn	Asn	Asn 260	Thr	Arg	Arg	Ser	1le 265	His	Ile	Gly	Pro	Gly 270	Arg	Ala
60	Phe	Tyr	Ala 275	Thr	Gly	Glu	Ile	11e 280	Gly	Asp	Ile	Arg	Gln 285	Ala	His	Cys
	Asn	Leu 290	Ser	Ser	Thr	Lys	Trp 295	Asn	Asn	Thr	Leu	Lys 300	Gln	lle	Val	Thr

```
Lys Leu Arg Glu His Phe Asn Lys Thr Ile Val Phe Asn His Ser Ser
       305
                           310
       Gly Gly Asp Pro Glu Ile Val Met His Ser Phe Asn Cys Gly Gly Glu
  5
       Phe Phe Tyr Cys Asn Thr Thr Pro Leu Phe Asn Ser Thr Trp Asn Tyr
       Thr Tyr Thr Trp Asn Asn Thr Glu Gly Ser Asn Asp Thr Gly Arg Asn
10
       Ile Thr Leu Gln Cys Arg Ile Lys Gln Ile Ile Asn Met Trp Gln Glu
15
       Val Gly Lys Ala Met Tyr Ala Pro Pro Ile Arg Gly Gln Ile Arg Cys
      Ser Ser Asn Ile Thr Gly Leu Leu Leu Thr Arg Asp Gly Gly Asn Asn
20
      Ser Glu Thr Glu Ile Phe Arg Pro Gly Gly Gly Asp Met Arg Asp Asn
25
      Trp Arg Ser Glu Leu Tyr Lys Tyr Lys Val Val Lys Ile Glu Pro Leu
      Gly Val Ala Pro Thr Lys Ala
          450
30
      (2) INFORMATION FOR SEQ ID NO:34:
           (i) SEQUENCE CHARACTERISTICS:
                (A) LENGTH: 37 base pairs
35
                (B) TYPE: Nucleic Acid
                (C) STRANDEDNESS: Single
                (D) TOPOLOGY: Linear
          (xi) SEQUENCE DESCRIPTION: SEQ ID NO:34:
40
      GGGAATTCGG ATCCAGAGCA GAAGACAGTG GCAATGA
      (2) INFORMATION FOR SEQ ID NO:35:
           (i) SEQUENCE CHARACTERISTICS:
                (A) LENGTH: 33 base pairs
45
                (B) TYPE: Nucleic Acid
                (C) STRANDEDNESS: Single
                (D) TOPOLOGY: Linear
          (xi) SEQUENCE DESCRIPTION: SEQ ID NO:35:
50
      CTCGAGCTCC TGAAGACAGT CAGACTCATC AAG
      (2) INFORMATION FOR SEQ ID NO:36:
55
           (i) SEQUENCE CHARACTERISTICS:
                (A) LENGTH: 39 base pairs
                (B) TYPE: Nucleic Acid
                (C) STRANDEDNESS: Single
                (D) TOPOLOGY: Linear
60
          (xi) SEQUENCE DESCRIPTION: SEQ ID NO:36:
     GGTCTAGAAG CTTTAGCCCA TAGTGCTTCC TGCTGCTCC
```

```
(2) INFORMATION FOR SEQ ID NO:37:
(i) SEQUENCE CHARACTERISTICS:
                 (A) LENGTH: 36 base pairs
                 (B) TYPE: Nucleic Acid
                 (C) STRANDEDNESS: Single (D) TOPOLOGY: Linear
 5
           (xi) SEQUENCE DESCRIPTION: SEQ ID NO:37:
      GGGCGGATCC TCGAGGTACC TGTRTGGAAA GAAGCA
10
      (2) INFORMATION FOR SEQ ID NO:38:
            (i) SEQUENCE CHARACTERISTICS:
                 (A) LENGTH: 38 base pairs (B) TYPE: Nucleic Acid
15
                 (C) STRANDEDNESS: Single
                 (D) TOPOLOGY: Linear
           (xi) SEQUENCE DESCRIPTION: SEQ ID NO:38:
      GGTCTAGAAG CTTTATGCTC CYAAGAACCC AAGGAACA 38
20
      (2) INFORMATION FOR SEQ ID NO:39:
            (i) SEQUENCE CHARACTERISTICS:
25
                 (A) LENGTH: 7 amino acids
                 (B) TYPE: Amino Acid
                 (C) STRANDEDNESS: Single
                 (D) TOPOLOGY: Linear
           (xi) SEQUENCE DESCRIPTION: SEQ ID NO:39:
30
           Ile Gly Pro Gly Arg Ala Phe
      (2) INFORMATION FOR SEQ ID NO:40:
35
           (i) SEQUENCE CHARACTERISTICS:
                 (A) LENGTH: 7 amino acids
                 (B) TYPE: Amino Acid
                 (C) STRANDEDNESS: Single
                 (D) TOPOLOGY: Linear
40
           (xi) SEQUENCE DESCRIPTION: SEQ ID NO:40:
           Ile Gly Pro Gly Arg Ala Trp
45
      (2) INFORMATION FOR SEQ ID NO:41:
           (i) SEQUENCE CHARACTERISTICS:
                 (A) LENGTH: 7 amino acids
                 (B) TYPE: Amino Acid
(C) STRANDEDNESS: Single
50
                 (D) TOPOLOGY: Linear
          (xi) SEQUENCE DESCRIPTION: SEQ ID NO:41:
           Leu Gly Pro Gly Ser Thr Phe
55
      (2) INFORMATION FOR SEQ ID NO:42:
            (i) SEQUENCE CHARACTERISTICS:
60
                 (A) LENGTH: 7 amino acids
                 (B) TYPE: Amino Acid
                 (C) STRANDEDNESS: Single
                 (D) TOPOLOGY: Linear
          (xi) SEQUENCE DESCRIPTION: SEQ ID NO:42:
65
```

```
Ile Gly Pro Gly Arg Val Leu 5
        (2) INFORMATION FOR SEQ ID NO:43:
  5
               (i) SEQUENCE CHARACTERISTICS:
                      (A) LENGTH: 7 amino acids
                     (B) TYPE: Amino Acid
(C) STRANDEDNESS: Single
(D) TOPOLOGY: Linear
10
             (xi) SEQUENCE DESCRIPTION: SEQ ID NO:43:
              Ile Gly Pro Gly Ser Ala Phe
15
        (2) INFORMATION FOR SEQ ID NO:44:
(i) SEQUENCE CHARACTERISTICS:
                     (A) LENGTH: 5 amino acids (B) TYPE: Amino Acid (C) STRANDEDNESS: Single
20
                     (D) TOPOLOGY: Linear
             (xi) SEQUENCE DESCRIPTION: SEQ 1D NO:44:
              Ile Gly Pro Gly Arg
25
```

WHAT IS CLAIMED IS:

5

30

1. An isolated polypeptide comprising an HIV gp120 amino acid sequence selected from the group consisting of Sequence ID Nos. 2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, and 28, and fragments thereof.

- The polypeptide of Claim 1 wherein the polypeptide
 additionally comprises a flag epitope sequence.
 - 3. The polypeptide of Claim 2 wherein the flag epitope sequence is HSV gD-1 flag epitope sequence.
- The polypeptide of Claim 2 wherein the flag epitope sequence is fused to the HIV gp120 amino acid sequence.
- 20 5. An oligonucleotide of not more than five kilobases encoding an HIV gpl20 polypeptide sequence comprising an amino acid sequence selected from the group consisting of Sequence ID Nos. 2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, and 28, and fragments thereof.
 - 6. The oligonucleotide of Claim 5 wherein the oligonucleotide includes a nucleotide sequence selected from the group consisting of Sequence ID Nos. 1, 3, 5, 7, 9, 11, 13, 15, 7, 19, 21, 23, 25, and 27, and fragments thereof.
- 7. The oligonucleotide of Claim 5 wherein the amino acid sequence encoded by the oligonucleotide
 35 additionally comprises a flag epitope.

8. The oligonucleotide of Claim 5 wherein the flag epitope is HSV gD-1 flag epitope.

- 9. The oligonucleotide of Claim 7 wherein the flag epitope is fused to the HIV gp120 amino acid sequence.
- 10. A vaccine comprising gp120 MN and an HIV gp120 polypeptide sequence selected from the group consisting of Sequence ID Nos. 2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, and 28, and fragments thereof in a suitable carrier.
 - 11. A vaccine comprising:

20

- 15 a. a first gp120 polypeptide sequence or a fragment thereof; and
 - b. a breakthrough isolate HIV gp120 polypeptide sequence or a fragment thereof from a vaccinee vaccinated with said first HIV gp120 polypeptide sequence;

wherein said HIV gp120 polypeptide sequences are in a suitable carrier.

- 12. The vaccine of Claim 11 wherein said first HIV
 25 gp120 polypeptide sequence comprises gp120 MN,
 gp120 A244, gp120 MN-GNE6 (Sequence ID No. 31), or
 gp120 MN-GNE8 (Sequence ID No. 33).
- 13. The vaccine of Claim 12 wherein said vaccine
 additionally comprises a second gp120 polypeptide
 sequence comprising gp120 MN, gp120 A244, gp120
 MN-GNE6 (Sequence ID No. 31), or gp120 MN-GNE8
 (Sequence ID No. 33), or a fragment thereof,
 wherein said second HIV gp120 polypeptide sequence
 is different from said first HIV gp120 polypeptide
 sequence.

14. The vaccine of Claim 13 wherein said first gp120 polypeptide sequence comprises gp120 MN and said second gp120 polypeptide sequence comprises gp120 A244.

5

10

15

20

- 15. The vaccine of Claim 14 wherein said breakthrough isolate comprises an HIV gp120 polypeptide sequence selected from the group consisting of Sequence ID Nos. 2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, and 28, and fragments thereof in a suitable carrier.
- 16. The vaccine of Claim 13 wherein said first gp120 polypeptide sequence comprises gp120 MN and said second gp120 polypeptide sequence comprises gp120 MN-GNE8 (Sequence ID No. 33).
- 17. The vaccine of Claim 16 wherein said breakthrough isolate HIV gp120 polypeptide sequence is an HIV gp120 polypeptide sequence selected from the group consisting of Sequence ID Nos. 2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, and 28, and fragments thereof, in a suitable carrier.
- 25 18. The vaccine of Claim 13 wherein said breakthrough isolate HIV gp120 polypeptide is from a vaccinee vaccinated with said first and second HIV gp120 polypeptide sequences.

19. A method for making an HIV vaccine comprising adding an HIV gp120 polypeptide sequence or fragments thereof from a breakthrough isolate from a vaccinee to the vaccine with which the vaccinee was vaccinated.

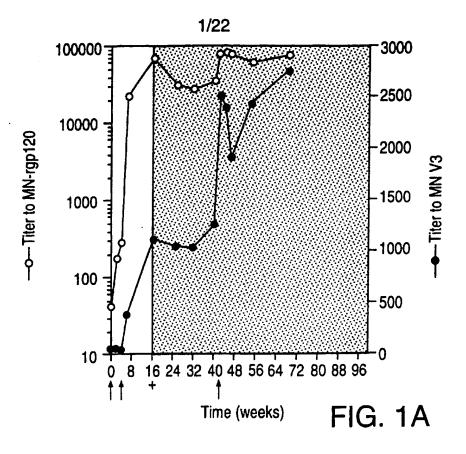
20. The vaccine of Claim 11 wherein said first gp120 polypeptide sequence is from a macrophage-tropic HIV-1 strain.

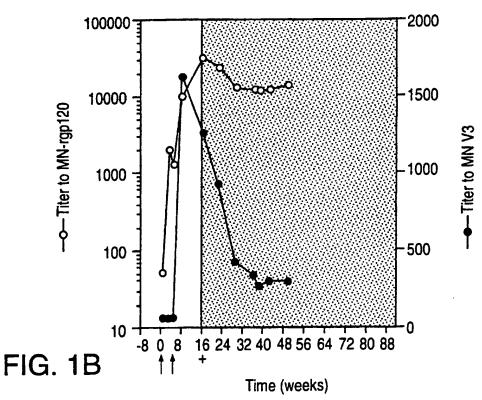
10

5

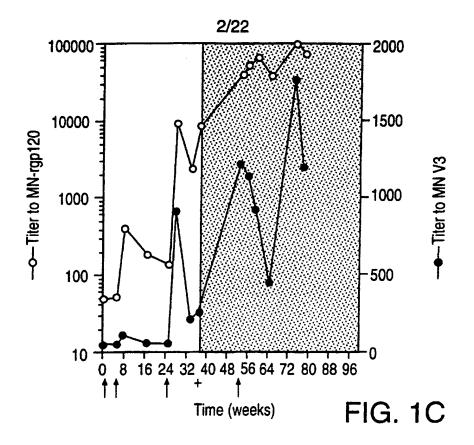
- 21. The vaccine of Claim 11 wherein said first gp120 polypeptide sequence is from a T-cell-tropic HIV-1 strain.
- 15 22. The vaccine of Claim 21 wherein said vaccine additionally comprises a second gp120 polypeptide sequence or a fragment, from a macrophage-tropic HIV-1 strain.
- 20 23. The vaccine of Claim 22 wherein said first and second gp120 polypeptide sequences bind to different chemokine receptors.
- 24. The vaccine of Claim 23 wherein said first gp120 polypeptide sequence binds to CC-CKR-5 and said second gp 120 polypeptide sequence binds to CXC-CKR-4.
- 25. The vaccine of Claim 11 wherein said vaccine additionally comprises an virus engineered to induce a cytotoxic T-cell response.

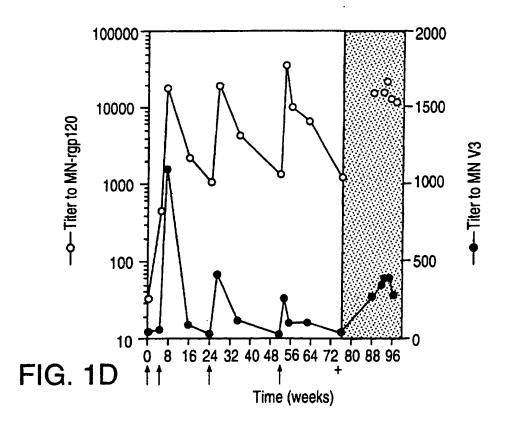
PCT/US97/09690

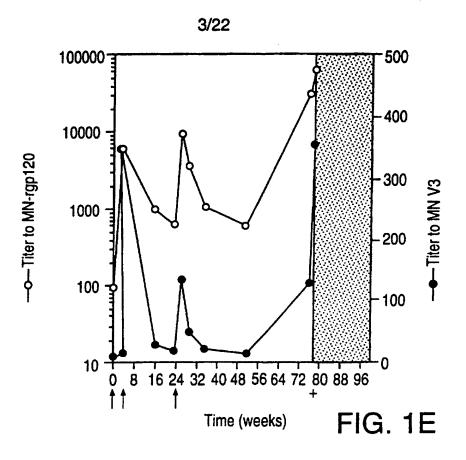


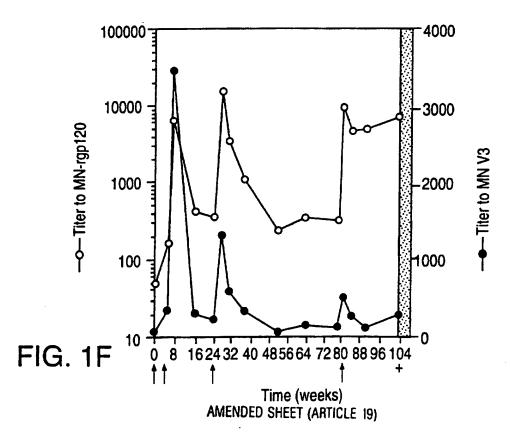


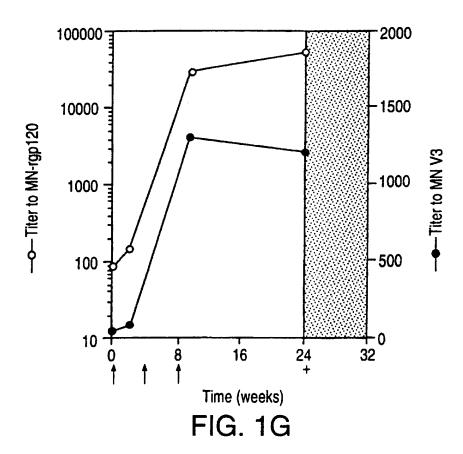
AMENDED SHEET (ARTICLE 19)



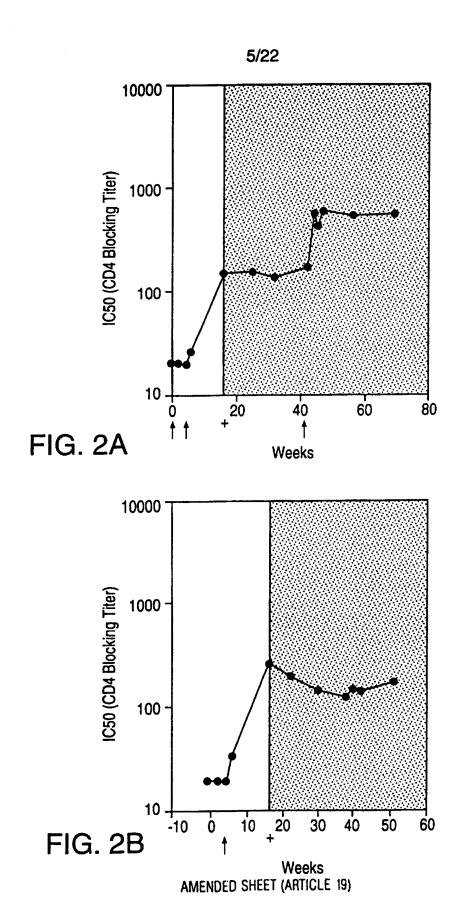


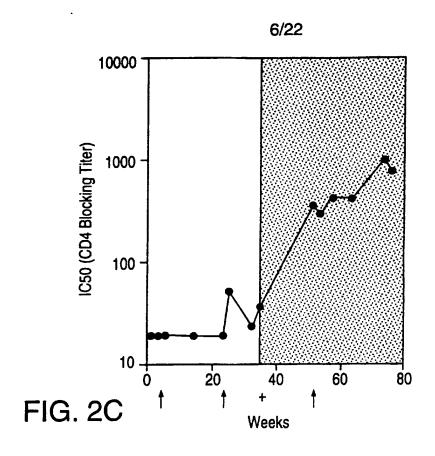


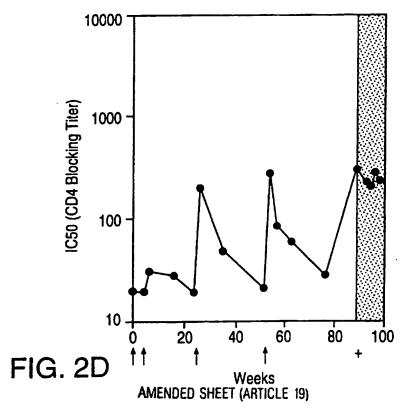


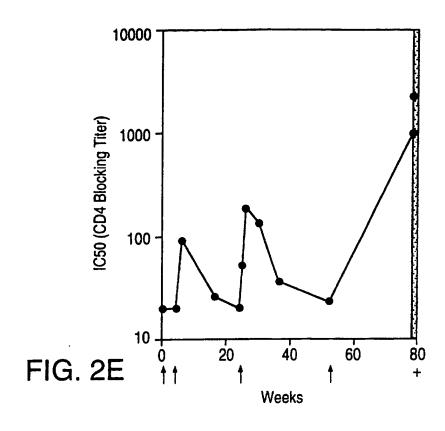


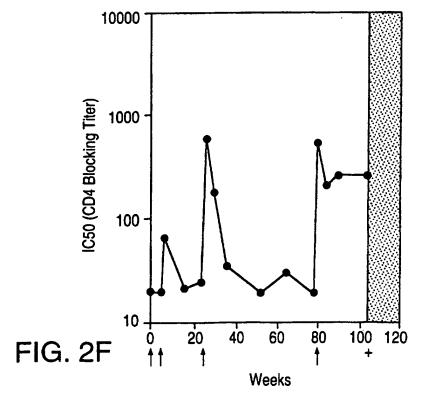
AMENDED SHEET (ARTICLE 19)





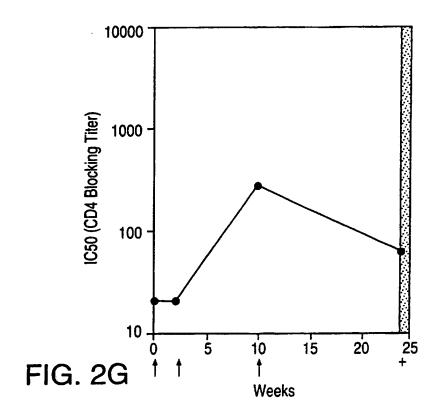






AMENDED SHEET (ARTICLE 19)

8/22



mmmmoomo>>oomm>>>>>> aaaaaaaaaaaaa ~~~~~~~~~~~~~~ **ZZZZZZ**ののZZZZZZZ 000000000000000 >>>>>>>>>> 0000000000000000 **444444444444 33333333333333 44444444444** 000000000000000 o*⋖⋖⋖⋖⋖⋖⋖⋖⋖⋖⋖⋖⋖* 00000000000000000 >>>>>>>>>>

C6.1 C6.5 C8.3 C8.6 C15.2 C7.2 C7.2 C7.10 C11.5 C10.5 C10.7 C10.7 C10.7

AMENDED SHEET (ARTICLE 19)

FIG. 3B		10/22	
VIEDFNAWKNIDMVEOMHEDIISLWDOSLKPCVKLTPLCITLNCTNWKKNDTKENFNAWKNIDMVEOMHEDIISLWDOSLKPCVKLTPLCITLNCTNWKKNDTKENFNAWKNIDMVEOMHEDIISLWDOSLKPCVKLTPLCITLNCTNWKKNDTKENFNAWKNNNNVEOMHEDIISLWDOSLKPCVKLTPLCVTLNCTIVWKENDTKTENFNAWKNNNNVEOMHEDIISLWDOSLKPCVKLTPLCVTLNCTDIINTENFNAWKNNNNVEOMHEDIISLWDOSLKPCVKLTPLCVTLNCTDIINTENFNAWKNNNNNVEOMHEDIISLWDOSLKPCVKLTPLCVTLNCTDIINTENFNAWKNNNNNVEOMHEDIISLWDOSLKPCVKLTPLCVTLNCSDIINTENFNAWKNNNNNVEOMHEDIISLWDOSLKPCVKLTPLCVTLNCSDIINTENFNAWKNNNNVEOMHEDIISLWDOSLKPCVKLTPLCVTLNCSDIINTENFNAWKNNNNNVEOMHEDIISLWDOSLKPCVKLTPLCVTLNCTDIINTENFNAWKNNNNNVEOMHEDIISLWDOSLKPCVKLTPLCVTLNCTDIINTENFNAWKNNNNVEOMHEDIISLWDOSLKPCVKLTPLCVTLNCTDIINTENFNAWKNNNNVEOMHEDIISLWDOSLKPCVKLTPLCVTLNCTDIINTENFNAWKNNNNVEOMHEDIISLWDOSLKPCVKLTPLCVTLNCTDIINTENFNAWKNNNNNVEOMHEDIISLWDOSLKPCVKLTPLCVTLNCTDIINTENFNAWKNNNNNVEOMHEDIISLWDOSLKPCVKLTPLCVTLNCTDIINTENFNAWKNNNNNVEOMHEDIISLWDOSLKPCVKLTPLCVTLNCTDIINTENFNAWKNNNNNVEOMHEDIISLWDOSLKPCVKLTPLCVTLNCTDIINTENFNAWKNNNNNVEOMHEDIISLWDOSLKPCVKLTPLCVTLNCTDIINTENFNAWKNNNNNVEOMHEDIISLWDOSLKPCVKLTPLCVTLNCTDIINTENFNAWKNNNNNVEOMHEDIISLWDOSLKPCVKLTPLCVTLNCTDIINTENFNAWKNNNNNVEOMHEDIISLWDOSLKPCVKLTPLCVTLNCTDIINTENFNAWKNNNNNVEOMHEDIISLWDOSLKPCVKLTPLCVTLNCTDIINTENFNAWKNNNNNVEOMHEDIISLWDOSLKPCVKLTPLCVTLNCTDIINTENFNAWKNNNNNVEOMHEDIISLWDOSLKPCVKLTPLCVTLNCTDIINTENFNAWKNNNNNVEOMHEDIISLWDOSLKPCVKLTPLCVTLNCTDIINTENFNAWKNNNNNVEOMHEDIISLWDOSLKPCVKLTPLCVTLNCTDIINTENFNAWKNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN	1	<u></u>	
VIEDFNMWKNDMVEOMHEDIISLWDOSLKPCVKLTPLCITLNCTNWKKNDMVEOMHEDIISLWDOSLKPCVKLTPLCITLNCTNWKKNDMVEOMHEDIISLWDOSLKPCVKLTPLCITLNCTNWKKNDWVEOMHEDIISLWDOSLKPCVKLTPLCITLNCTNWKKNNOWYEOMHEDIISLWDOSLKPCVKLTPLCVTLNCTNWKNNOWYEOMHEDIISLWDOSLKPCVKLTPLCVTLNCTDIVTENFNMWKNNNWVEOMHEDIISLWDOSLKPCVKLTPLCVTLNCTDIVTENFNMWKNNNMVEOMHEDIISLWDOSLKPCVKLTPLCVTLNCTDIVTENFNMWKNNNMVEOMHEDIISLWDOSLKPCVKLTPLCVTLNCSDIVTENFNMWKNNNMVEOMHEDIISLWDOSLKPCVKLTPLCVTLNCSDIVTENFNMWKNNNMVEOMHEDIISLWDOSLKPCVKLTPLCVTLNCSDIVTENFNMWKNNNMVEOMHEDIISLWDOSLKPCVKLTPLCVTLNCSDIVTENFNMWKNNNMVEOMHEDIISLWDOSLKPCVKLTPLCVTLNCTDIVTENFNMWKNNNMVEOMHEDIISLWDOSLKPCVKLTPLCVTLNCTDIVTENFNMWKNNNMVEOMHEDIISLWDOSLKPCVKLTPLCVTLNCTDIVTENFNMWKNNNMVEOMHEDIISLWDOSLKPCVKLTPLCVTLNCTDIVTENFNMWKNNNMVEOMHEDIISLWDOSLKPCVKLTPLCVTLNCTDIVTENFNMWKNNNMVEOMHEDIISLWDOSLKPCVKLTPLCVTLNCTDIVTENFNMWKNNNMVEOMHEDIISLWDOSLKPCVKLTPLCVTLNCTDIVTENFNMWKNNNMVEOMHEDIISLWDOSLKPCVKLTPLCVTLNCTDIVTENFNMWKNNNMVEOMHEDIISLWDOSLKPCVKLTPLCVTLNCTDIVTENFNMWKNNNMVEOMHEDIISLWDOSLKPCVKLTPLCVTLNCTDIVTENFNMWKNNNMVEOMHEDIISLWDOSLKPCVKLTPLCVTLNCTDIVTENFNMWKNNNMVEOMHEDIISLWDOSLKPCVKLTPLCVTLNCTDIVTENFNMWKNNNMVEOMHEDIISLWDOSLKPCVKLTPLCVTLNCTDIVTENFNMWKNNNMVEOMHEDIISLWDOSLKPCVKLTPLCVTLNCTDIVTENFNMWKNNNMVEOMHEDIISLWDOSLKPCVKLTPLCVTLNCTDIVTENFNMWKNNNMVEOMHEDIISLWDOSLKPCVKLTPLCVTLNCTDIVTENFNMWKNNNMVEOMHEDIISLWDOSLKPCVKLTPLCVTLNCTDIVTENFNMWKNNNMVEOMHEDIISLWDOSLKPCVKLTPLCVTLNCTDIVTENFNMWKNNNNWVEOMHEDIISLWDOSLKPCVKLTPLCVTLNCTDIVTENFNMWKNNNNWVEOMHEDIISLWDOSLKPCVKLTPLCVTLNCTDIVTENFNMWKNNNNWVEOMHEDIISLWDOSLKPCVKLTPLCVTLNCTDIVTENFNMWKNNNNWVEOMHEDIISLWDOSLKPCVKLTPLCVTLNCTDIVTENFNMWKNNNNWVEOMHEDIISLWDOSLKPCVKLTPLCVTLNCTDIVTENFNMWKNNNNWVEOMHEDIISLWDOSLKPCVKLTPLCVTLNCTDIVTENFNMWKNNNNNUMVEOMHEDIISLWDOSLKPCVKLTPLCVTLNCTDIVTENFNMWKNNNNNNUMVEOMHEDIISLWDOSLKPCVKLTPLCVTLNCTDIVTENFNMWKNNNNNNNUMVEOMHEDIISLWCONTONTONTONTONTONTONTONTONTONTONTONTONTO		\times ×····································	
VIEDFNMWKNUDMVEOMHEDIISLWDOSLKPCVKLTPLCIITLNCTNWKKNVTENFNMWKNNDMVEOMHEDIISLWDOSLKPCVKLTPLCIITLNCTNWKKNVTENFNMWKNNNMVEOMHEDIISLWDOSLKPCVKLTPLCIITLNCTNWKKNNWWKNNNMVEOMHEDIISLWDOSLKPCVKLTPLCVTLNCTI VIENFNMWKNNNMVEOMHEDIISLWDOSLKPCVKLTPLCVTLNCTI VIENFNMWKNNNMVEOMHEDIISLWDOSLKPCVKLTPLCVTLNCTD VIENFNMWKNNMVEOMHEDIISLWDOSLKPCVKLTPLCVTLNCSD VIENFNMWKNNMVEOMHEDIISLWDOSLKPCVKLTPLCVTLNCSD VIENFNMWKNNMVEOMHEDIISLWDOSLKPCVKLTPLCVTLNCSD VIENFNMWKNNMVEOMHEDIISLWDOSLKPCVKLTPLCVTLNCSD VIENFNMWKNNNMVEOMHEDIISLWDOSLKPCVKLTPLCVTLNCSD VIENFNMWKNNNMVEOMHEDIISLWDOSLKPCVKLTPLCVTLNCTD VIENFNMWKNNNMVEOMHEDIISLMMUMUM VEOMHEDIISLMMUM VEOMHEDIISLMMUM VEOMHEDIISLMMUM VEOMHEDIISLMMUM VEOMHEDIISLMMUM VEOMHEDIISLMMUM VEOMHEDIISLMMUM VEOMHEDIISLMMUM VEOMHEDIISLMMUM VEOMHEDIIN VEOMHEDIISLMMUM VEOMHEDIISLMMUM VEOMHEDIIN VE	1	⊢ ⊢ · · · · · · · · ZZ · · · · · ·	
VTEDFNMWKNNDMVEQMHEDIISLWDQSLKPCVKLTPLCIITLNCTNWKKNTENFNMWKNNNWVEQMHEDIISLWDQSLKPCVKLTPLCIITLNCTNWKEVTENFNMWKNNNMVEQMHEDIISLWDQSLKPCVKLTPLCVTLNCTNWKEVTENFNMWKNNNMVEQMHEDIISLWDQSLKPCVKLTPLCVTLNCTNWKEVTENFNMWKNNNMVEQMHEDIISLWDQSLKPCVKLTPLCVTLNCTDIVTENFNMWKNNNMVEQMHEDIISLWDQSLKPCVKLTPLCVTLNCTDIVTENFNMWKNNNMVEQMHEDIISLWDQSLKPCVKLTPLCVTLNCSDIVTENFNMWKNNNMVEQMHEDIISLWDGSLKPCVKLTPLCVTLNCSDIVTENFNMWKNNNMVEQMHEDIISLWDGSLKPCVKLTPLCVTLNCSDIVTENFNMWKNNNMVEQMHEDIISLWDGSLKPCVKLTPLCVTLNCSDIVTENFNMWKNNNMVEQMHEDIISLWDQSLKPCVKLTPLCVTLNCSDIVTENFNMWKNNNMVEQMHEDIISLWDQSLKPCVKLTPLCVTLNCTDIVTENFNMWKNNNMVEQMHEDIISLWDQSLKPCVKLTPLCVTLNCTDIVTENFNNWKNNNMVEQMHEDIISLWDQSLKPCVKLTPLCVTLNCTDIVTENFNNWKNNNMVEQMHEDIISLWDQSLKPCVKLTPLCVTLNCTDIVTENFNNWKNNNMVEQMHEDIISLWDQSLKPCVKLTPLCVTLNCTDIVTENFNNWKNNNMVEQMHEDIISLWDQSLKPCVKLTPLCVTLNCTDIVTENFNNWKNNNMVEQMHEDIISLWDQSLKPCVKLTPLCVTLNCTDIVTENFNNWKNNNMVEQMHEDIISLWDQSLKPCVKLTPLCVTLNCTDIVTENFNNWKNNNMVEQMHEDIISLWDQSLKPCVKLTPLCVTLNCTDIVTENFNNWKNNNMVEQMHEDIISLWDQSLKPCVKLTPLCVTLNCTDIVTENFNNWKNNNMVEQMHEDIISLWDQSLKPCVKLTPLCVTLNCTDIVTENFNNWKNNNMVEQMHEDIISLWDQSLKPCVKLTPLCVTLNCTDIVTENFNNWKNNNMVEQMHEDIINNUMVEQNHEQDIISLWDQSLKPCVKLTPLCVTLNCTDIVTENFNNWKNNNMVEQMHEDIINNUMVEQNHEQDIISLWDQSLKPCVKLTPLCVTLNCTDIVTENFNNWKNNNMVEQMHEDIINNUMVEQNHEQDIISLWDQSLKPCVKLTPLCVTLNCTDIVTENFNNWKNNNMVEQMHEQDIISLWDQSLKPCVKLTPLCVTLNCTDIVTENFNNMWKNNNMVEQNHEQDIISLWDQSLKPCVKLTPLCVTLNCTDIVTENFNNMWKNNNMVEQNHEQDIISLWDQSLKPCVKLTPLCVTLNCTDIVTENFNNMWKNNNMVEQNHEQDIISLWDQSLKPCVKLTPLCVTLNCTDIVTENFNMWKNNNMVEQNHEQDIISLWDQSLKPCVKLTPLCVTLNCTDIVTENFNMWKNNNMVEQNHEQDIISLWDQSLKPCVKLTPLCVTLNCTDIVTENFNMWKNNNMVEQNHEQDIISLWDQSLKPCVKLTPLCVTLNCTDIVTENFNMWKNNNMVEQNHEQDIISLWDQSLKPCVKLTPLCVTLNCTDIVTENFNMWKNNNMVEQNHEQDIISLWDQSLKPCVKLTPLCVTLNCTDIVTENCTDIVTENTFNMWKNNNMVEQNHEQDIISLWDQSLKPCVKLTPLCVTLNCTDIVTENTFNMMVEQNHEQDIISLWDQSLKPCVKLTPLCVTLNCTDIVTENTFNMMVEQNHEQDIISLWDQSLKPCVKLTPLCVTLNCTDIVTENTFNMMVEQNHEQDIISLWDQSLKPCVKLTPLCVTLNCTDIVTENTFNMMVEQNHEQDIISLMVEQUARTENTFNMMVEQNATURCTURCTURCTURCTURCTURCTURCTURCTURCTURC	1		
VTEDFNMWKNNDMVEQMHEDIISLWDQSLKPCVKLTPLCITLNCTNWK VTENFNMWKNNDMVEQMHEDIISLWDQSLKPCVKLTPLCITLNCTNWK VTENFNMWKNNNMVEQMHEDIISLWDQSLKPCVKLTPLCITLNCTNWK VTENFNMWKNNNMVEQMHEDIISLWDQSLKPCVKLTPLCVTLNCTD VTENFNMWKNNNMVEQMHEDIISLWDQSLKPCVKLTPLCVTLNCTD VTENFNMWKNNNMVEQMHEDIISLWDQSLKPCVKLTPLCVTLNCTD VTENFNMWKNNNMVEQMHEDIISLWDQSLKPCVKLTPLCVTLNCSD VTENFNMWKNNNMVEQMHEDIISLWDQSLKPCVKLTPLCVTLNCSD VTENFNMWKNNNMVEQMHEDIISLWDQSLKPCVKLTPLCVTLNCSD VTENFNMWKNNNMVEQMHEDIISLWDQSLKPCVKLTPLCVTLNCSD VTENFNMWKNNNMVEQMHEDIISLWDQSLKPCVKLTPLCVTLNCTD VTENFNMWKNNMVEQMHEDIISLWDQSLKPCVKLTPLCVTLNCTD VTENFNMWKNNMVEQMHEDIISLWDQSLKPCVKLTPLCVTLNCTD VTENFNMWKNNMVEQMHEDIISLWDQSLKPCVKLTPLCVTLNCTD VTENFNMWKNNMVEQMHEDIISLWDQSLKPCVKLTPLCVTLNCTD VTENFNMWKNNMVEQMHEDIISLWDQSLKPCVKLTPLCVTLNCTD VTENFNMWKNNMVEQMHEDIISLWCCONTUNCTD	>	ZZ · · · · · · >> · · · · ·	
VTEDFNMWKNDMVEQMHEDIISLWDQSLKPCVKLTPLCIITLNCTNW VTENFNMWKNNDMVEQMHEDIISLWDQSLKPCVKLTPLCIITLNCTNW VTENFNMWKNNNMVEQMHEDIISLWDQSLKPCVKLTPLCVTLNCTNW VTENFNMWKNNNMVEQMHEDIISLWDQSLKPCVKLTPLCVTLNCTD VTENFNMWKNNNMVEQMHEDIISLWDQSLKPCVKLTPLCVTLNCTD VTENFNMWKNNNMVEQMHEDIISLWDQSLKPCVKLTPLCVTLNCTD VTENFNMWKNNNMVEQMHEDIISLWDGSLKPCVKLTPLCVTLNCSD VTENFNMWKNNNMVEQMHEDIISLWDGSLKPCVKLTPLCVTLNCSD VTENFNMWKNNNMVEQMHEDIISLWDGSLKPCVKLTPLCVTLNCSD VTENFNMWKNNNMVEQMHEDIISLWDGSLKPCVKLTPLCVTLNCSD VTENFNMWKNNNMVEQMHEDIISLWDGSLKPCVKLTPLCVTLNCTD VTENFNMWKNNNMVEQMHEDIISLWDQSLKPCVKLTPLCVTLNCTD VTENFNMWKNNNMVEQMHEDIISLWDQSLKPCVKLTPLCVTLNCTD VTENFNMWKNNNMVEQMHEDIISLWDQSLKPCVKLTPLCVTLNCTD VTENFNMWKNNNMVEQMHEDIISLWDQSLKPCVKLTPLCVTLNCTD VTENFNMWKNNNMVEQMHEDIISLWDQSLKPCVKLTPLCVTLNCTD VTENFNMWKNNNMVEQMHEDIISLWDQSLKPCVKLTPLCVTLNCTD VTENFNMWKNNNMVEQMHEDIISLWDQSLKPCVKLTPLCVTLNCTD	1	⊼ m · · · · · · · · · · · · · ·	
VTEDFNMWKNNDMVEQMHEDIISLWDQSLKPCVKLTPLCITLNCTNVTEDFNMWKNNDMVEQMHEDIISLWDQSLKPCVKLTPLCITLNCTNVTENFNMWKNNNWVEQMHEDIISLWDQSLKPCVKLTPLCVTLNCTNVTENFNMWKNNNWVEQMHEDIISLWDQSLKPCVKLTPLCVTLNCTDVTENFNMWKNNNWVEQMHEDIISLWDQSLKPCVKLTPLCVTLNCTDVTENFNMWKNNNWVEQMHEDIISLWDQSLKPCVKLTPLCVTLNCTDVTENFNMWKNNMWVEQMHEDIISLWDQSLKPCVKLTPLCVTLNCSDVTENFNMWKNNMWVEQMHEDIISLWDGSLKPCVKLTPLCVTLNCSDVTENFNMWKNNMWVEQMHEDIISLWDGSLKPCVKLTPLCVTLNCSDVTENFNMWKNNMWVEQMHEDIISLWDQSLKPCVKLTPLCVTLNCTDVTENFNMWKNNMWVEQMHEDIISLWDQSLKPCVKLTPLCVTLNCTDVTENFNMWKNNMWVEQMHEDIISLWDQSLKPCVKLTPLCVTLNCTDVTENFNMWKNNMWVEQMHEDIISLWDQSLKPCVKLTPLCVTLNCTDVTENFNMWKNNMWVEQMHEDIISLWDQSLKPCVKLTPLCVTLNCTDVTENFNMWKNNMWVEQMHEDIISLWDQSLKPCVKLTPLCVTLNCTDVTENFNMWKNNMWVEQMHEDIISLWDQSLKPCVKLTPLCVTLNCTDVTENFNMWKNNMWVEQMHEDIISLWDQSLKPCVKLTPLCVTLNCTDVTGTDVTGTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	-	$\mathbf{x}\mathbf{x}$ · · · · · · · · · · · · · ·	
VTEDFNMWKNDDMVEOMHEDIISLWDOSLKPCVKLTPLCITLNOTTENFNMWKNNDMVEOMHEDIISLWDOSLKPCVKLTPLCITLNOTTENFNMWKNNNWVEOMHEDIISLWDOSLKPCVKLTPLCITLNOTTVTENFNMWKNNNWVEOMHEDIISLWDOSLKPCVKLTPLCVTLNOTTVTENFNMWKNNNWVEOMHEDIISLWDOSLKPCVKLTPLCVTLNOTTENFNMWKNNNWVEOMHEDIISLWDOSLKPCVKLTPLCVTLNOTTENFNMWKNNNWVEOMHEDIISLWDOSLKPCVKLTPLCVTLNOSVTENFNMWKNNNWVEOMHEDIISLWDOSLKPCVKLTPLCVTLNOSVTENFNMWKNNNWVEOMHEDIISLWDOSLKPCVKLTPLCVTLNOSVTENFNMWKNNNWVEOMHEDIISLWDOSLKPCVKLTPLCVTLNOSVTENFNMWKNNNWVEOMHEDIISLWDOSLKPCVKLTPLCVTLNOTTENFNMWKNNMVEOMHEDIISLMOTTENFNMWKNNMVEOMHEDIISLMOTTENFNMWKNNMVEOMHEDIISLMOTTENFNMMKNNMVEOMHEDIISLMOTTENFNMMKNNMVEOMHEDIISLMOTTENFNMMKNNMVEOMHEDIISLMOTTENFNMMVEOMHEDIISLMOTTENFNMMVEOMHEDIISLMOTTENFNMMKNNMVEOMHEDIISLMOTTENFNMMVEOMHEDIISLMOTTENFNMMVEOMHEDIISLMOTTENFNMMVEOMHEDIISLMOTTENFNMMVEOMHEDIISLMOTTENFNMMVEOMHEDIISLMOTTENFNMMVEOMHEDIISLMOTTENFNM		≽ ≱ · · <u>· · · · · · · · · · · · · · · · · </u>	
C1- VTEDFNMWKNDMVEQMHED11SLWDQSLKPCVKLTPLC1TLNC VTENFNMWKNNMVEQMHED11SLWDQSLKPCVKLTPLC1TLNC VTENFNMWKNNMVEQMHED11SLWDQSLKPCVKLTPLC1TLNC VTENFNMWKNNMVEQMHED11SLWDQSLKPCVKLTPLCVTLNC VTENFNMWKNNMVEQMHED11SLWDQSLKPCVKLTPLCVTLNC VTENFNMWKNNMVEQMHED11SLWDQSLKPCVKLTPLCVTLNC VTENFNMWKNNMVEQMHED11SLWDGSLKPCVKLTPLCVTLNC VTENFNMWKNNMVEQMHED11SLWDGSLKPCVKLTPLCVTLNC VTENFNMWKNNMVEQMHED11SLWDGSLKPCVKLTPLCVTLNC VTENFNMWKNNMVEQMHED11SLWDGSLKPCVKLTPLCVTLNC VTENFNMWKNNMVEQMHED11SLWDGSLKPCVKLTPLCVTLNC VTENFNMWKNNMVEQMHED11SLWDQSLKPCVKLTPLCVTLNC VTENFNMWKNNNMVEQMHED11SLWDQSLKPCVKLTPLCVTLNC VTENFNMWKNNMVEQMHED11SLWDQSLKPCVKLTPLCVTLNC VTENFNMWKNNMVEQMHED11SLWDQSLKPCVKLTPLCVTLNC VTENFNMWKNNMVEQMHED11SLWDQSCKCOVKLTPLCVTLNC VTENFNMWKNNMVEQMHED11SLWDQSCKCOVKLTPLCVTLNC VTENFNMWKNNMVEQMHED11SCMMHED11SCMMTPLCVTLNC VTENFNMWKNNMVEQMHED11SCMMTPLCVTLNC VTENFNMWKNNMVEQMHED11SCMMTPLCVTLNC VTENFNMWKNNMVEQMHED11SCMMTPLCVTLNC VTENFNMWKNNMVEQMHED11SCMMTPLCVTLNC VTENFNMWKNNMVCNC VTENFNMMKNNMVEQMMHED11SCMMTPLCVTLNC VTENFNMMKNNMVCNC VTENFNMMKNNMVCNC VTENFNMMKNNMVCNC VTENFNMMKNNMMCCONTCN			
VTEDFNMWKNDMVEQMHEDIISLWDQSLKPCVKLTPLCITLNVTEDFNMWKNNDMVEQMHEDIISLWDQSLKPCVKLTPLCITLNVTENFNMWKNNMVEQMHEDIISLWDQSLKPCVKLTPLCVTLNVTENFNMWKNNMVEQMHEDIISLWDQSLKPCVKLTPLCVTLNVTENFNMWKNNMVEQMHEDIISLWDQSLKPCVKLTPLCVTLNVTEDFNMWKNNMVEQMHEDIISLWDQSLKPCVKLTPLCVTLNVTEDFNMWKNNMVEQMHEDIISLWDQSLKPCVKLTPLCVTLNVTEDFNMWKNNMVEQMHEDIISLWDQSLKPCVKLTPLCVTLNVTENFNMWKNNMVEQMHEDIISLWDQSLKPCVKLTPLCVTLNVTTLNTNTNTNTNTNTNTNTNTNTNTNTNTNTNTNT	١		
VIEDFNMWKNDMVEQMHEDIISLWDQSLKPCVKLTPLGITL VIENFNMWKNNMNVEQMHEDIISLWDQSLKPCVKLTPLCITL VIENFNMWKNNMNVEQMHEDIISLWDQSLKPCVKLTPLCLTL VIENFNMWKNNMNVEQMHEDIISLWDQSLKPCVKLTPLCVTL VIENFNMWKNNMNVEQMHEDIISLWDQSLKPCVKLTPLCVTL VIENFNMWKNNMNVEQMHEDIISLWDQSLKPCVKLTPLCVTL VIEDFNMWKNNMNVEQMHEDIISLWDQSLKPCVKLTPLCVTL VIEDFNMWKNNMNVEQMHEDIISLWDQSLKPCVKLTPLCVTL VIENFNMWKNNMNVEQMHEDIISLWDQSLKPCVKLTPLCVTL	Ţ	· · · · · · · · · · · · · · · · · · ·	
C1 VTEDFNMWKNUDMVEQMHEDIISLWDQSLKPCVKLTPLC VTENFNMWKNNMWVEQMHEDIISLWDQSLKPCVKLTPLC VTENFNMWKNNMWVEQMHEDIISLWDQSLKPCVKLTPLC VTENFNMWKNNMWVEQMHEDIISLWDQSLKPCVKLTPLC VTENFNMWKNNMWVEQMHEDIISLWDQSLKPCVKLTPLC VTENFNMWKNNMWVEQMHEDIISLWDQSLKPCVKLTPLC VTENFNMWKNNMWVEQMHEDIISLWDGSLKPCVKLTPLC VTENFNMWKNNMWVEQMHEDIISLWDGSLKPCVKLTPLC VTENFNMWKNNMWVEQMHEDIISLWDGSLKPCVKLTPLC VTENFNMWKNNMWVEQMHEDIISLWDGSLKPCVKLTPLC VTENFNMWKNNMVEQMHGDIISLWDQSLKPCVKLTPLC VTENFNMWKNNMWVEQMHGDIISLWDQSLKPCVKLTPLC		ZZZZZZZZZZZZZZZ	1
C1 VTEDFNWWKNUDMVEQMHEDIISLWDQSLKPCVKLTPLC VTENFNMWKNNMWVEQMHEDIISLWDQSLKPCVKLTPLC VTENFNMWKNNMWVEQMHEDIISLWDQSLKPCVKLTPLC VTENFNMWKNNMWVEQMHEDIISLWDQSLKPCVKLTPLC VTENFNMWKNNMWVEQMHEDIISLWDQSLKPCVKLTPLC VTENFNMWKNNMWVEQMHEDIISLWDGSLKPCVKLTPLC VTENFNMWKNNMWVEQMHEDIISLWDGSLKPCVKLTPLC VTENFNMWKNNMWVEQMHEDIISLWDGSLKPCVKLTPLC VTENFNMWKNNMWVEQMHEDIISLWDGSLKPCVKLTPLC VTENFNMWKNNMWVEQMHEDIISLWDGSLKPCVKLTPLC VTENFNMWKNNMVEQMHEDIISLWDGSLKPCVKLTPLC VTENFNMWKNNMVEQMHEDIISLWDQSLKPCVKLTPLC VTENFNMWKNNMWVEQMHEDIISLWDQSLKPCVKLTPLC VTENFNMWKNNMWVEQMHEDIISLWDQSLKPCVKLTPLC VTENFNMWKNNMWVEQMHEDIISLWDQSLKPCVKLTPLC VTENFNMWKNNMWVEQMHEDIISLWDQSLKPCVKLTPLC VTENFNMWKNNMWFNNMVEQMHEDIISLWDQSLKPCVKLTPLC VTENFNMWKNNMVEQMHEDIISLMDGSLKPCVKLTPLC VTENFNMWMHEDIISLMDGSLKPCVKLTPLC	-		
C1 VTEDFNMWKNUDMVEQMHEDIISLWDQSLKPCVKLTPLC VTENFNMWKNNMWVEQMHEDIISLWDQSLKPCVKLTPLC VTENFNMWKNNMWVEQMHEDIISLWDQSLKPCVKLTPLC VTENFNMWKNNMWVEQMHEDIISLWDQSLKPCVKLTPLC VTENFNMWKNNMWVEQMHEDIISLWDQSLKPCVKLTPLC VTENFNMWKNNMWVEQMHEDIISLWDQSLKPCVKLTPLC VTENFNMWKNNMWVEQMHEDIISLWDGSLKPCVKLTPLC VTENFNMWKNNMWVEQMHEDIISLWDGSLKPCVKLTPLC VTENFNMWKNNMWVEQMHEDIISLWDGSLKPCVKLTPLC VTENFNMWKNNMWVEQMHEDIISLWDGSLKPCVKLTPLC VTENFNMWKNNMVEQMHGDIISLWDQSLKPCVKLTPLC VTENFNMWKNNMWVEQMHGDIISLWDQSLKPCVKLTPLC	-		
C1 VTEDFNMWKNDMVEQMHED!ISLWDQSLKPCVKLTPL VTENFNMWKNNMVEQMHED!ISLWDQSLKPCVKLTPL VTENFNMWKNNMVEQMHED!ISLWDQSLKPCVKLTPL VTENFNMWKNNMVEQMHED!ISLWDQSLKPCVKLTPL VTENFNMWKNNMVEQMHED!ISLWDQSLKPCVKLTPL VTENFNMWKNNMVEQMHED!ISLWDQSLKPCVKLTPL VTENFNMWKNNMVEQMHED!ISLWDQSLKPCVKLTPL VTENFNMWKNNMVEQMHED!ISLWDGSLKPCVKLTPL VTENFNMWKNNMVEQMHED!ISLWDGSLKPCVKLTPL VTENFNMWKNNMVEQMHED!INLWDQSLKPCVKLTPL VTENFNMWKNNMVEQMHED!INLWDQSLKPCVKLTPL VTENFNMWKNNMVEQMHED!INLWDQSLKPCVKLTPL VTENFNMWKNNMVEQMHED!INLWDQSLKPCVKLTPL VTENFNMWKNNMVEQMHED!ISLWDQSLKPCVKLTPL VTENFNMWKNNMVEQMHED!ISLWDQSLKPCVKLTPL VTENFNMWKNNMVEQMHED!ISLWDQSLKPCVKLTPL VTENFNMWKNNMVEQMHED!ISLWDQSLKPCVKLTPL VTENFNMWKNNMVEQMHED!ISLWDQSLKPCVKLTPL VTENFNMWKNNMVEQMHED!ISLWDQSLKPCVKLTPL VTENFNMWKNNNMVEQMHED!ISLWDQSLKPCVKLTPL VTENFNMWKNNNMVEQMHED!ISLWDQSLKPCVKLTPL VTENFNMWKNNNMVEQMHED!ISLWDQSLKPCVKLTPL VTENFNMWKNNNMVEQMHED!ISLWDQSLKPCVKLTPL			
C1- VTEDFNMWKNDMVEQMHED11SLWDQSLKPCVKLT VTENFNMWKNNMWVEQMHED11SLWDQSLKPCVKLT VTENFNMWKNNMWVEQMHED11SLWDQSLKPCVKLT VTENFNMWKNNMWVEQMHED11SLWDQSLKPCVKLT VTENFNMWKNNMWVEQMHED11SLWDQSLKPCVKLT VTENFNMWKNNMWVEQMHED11SLWDQSLKPCVKLT VTEDFNMWKNNMWVEQMHED11SLWDQSLKPCVKLT VTEDFNMWKNNMWVEQMHED11SLWDQSLKPCVKLT VTENFNMWKNNMWVEQMHED11SLWDQSLKPCVKLT VTENFNMWKNNMWVEQMHED11SLWDQSLKPCVKLT VTENFNMWKNNMWVEQMHED11SLWDQSLKPCVKLT VTENFNMWKNNMWVEQMHED11SLWDQSLKPCVKLT VTENFNMWKNNMWVEQMHED11SLWDQSLKPCVKLT VTENFNMWKNNMVEQMHED11SLWDQSLKPCVKLT VTENFNMWKNNMWVEQMHED11SLWDQSLKPCVKLT VTENFNMWKNNMWKNNMVEQMHED11SLWDQSLKPCVKLT VTENFNMWKNNMWKNNMVEQMHED11SLWDQSLKPCVKLT VTENFNMWKNNMWKNNMWVEQMHED11SLWDQSLKPCVKLT VTENFNMWKNNMWKNNMWVEQMHED11SLWDQSLKPCVKLT VTENFNMWKNNMWKNNMWMVEQMHED11SLWDQSLKPCVKLT VTENFNMWKNNMWKNNMWNGCONNMVEQNHED11SLWDQSLKPCVKLT VTENFNMWKNNMWKNNMWKNNMWKNNMMVEQNHED11SLWDQSLKF		000000000000000000000000000000000000000	
C1- VTEDFNMWKNDMVEQMHED11SLWDQSLKPCVKLT VTENFNMWKNNMWVEQMHED11SLWDQSLKPCVKLT VTENFNMWKNNMWVEQMHED11SLWDQSLKPCVKLT VTENFNMWKNNMWVEQMHED11SLWDQSLKPCVKLT VTENFNMWKNNMWVEQMHED11SLWDQSLKPCVKLT VTENFNMWKNNMWVEQMHED11SLWDQSLKPCVKLT VTEDFNMWKNNMWVEQMHED11SLWDQSLKPCVKLT VTEDFNMWKNNMWVEQMHED11SLWDQSLKPCVKLT VTENFNMWKNNMWVEQMHED11SLWDQSLKPCVKLT VTENFNMWKNNMWVEQMHED11SLWDQSLKPCVKLT VTENFNMWKNNMWVEQMHED11SLWDQSLKPCVKLT VTENFNMWKNNMWVEQMHED11SLWDQSLKPCVKLT VTENFNMWKNNMWVEQMHED11SLWDQSLKPCVKLT VTENFNMWKNNMVEQMHED11SLWDQSLKPCVKLT VTENFNMWKNNMWVEQMHED11SLWDQSLKPCVKLT VTENFNMWKNNMWKNNMVEQMHED11SLWDQSLKPCVKLT VTENFNMWKNNMWKNNMVEQMHED11SLWDQSLKPCVKLT VTENFNMWKNNMWKNNMWVEQMHED11SLWDQSLKPCVKLT VTENFNMWKNNMWKNNMWVEQMHED11SLWDQSLKPCVKLT VTENFNMWKNNMWKNNMWMVEQMHED11SLWDQSLKPCVKLT VTENFNMWKNNMWKNNMWNGCONNMVEQNHED11SLWDQSLKPCVKLT VTENFNMWKNNMWKNNMWKNNMWKNNMMVEQNHED11SLWDQSLKF			İ
C1 VTEDFNMWKNDMVEQMHEDIISLWDDSLKPCVK VTENFNMWKNNMVEQMHEDIISLWDDSLKPCVK VTENFNMWKNNNMVEQMHEDIISLWDDSLKPCVK VTENFNMWKNNMVEQMHEDIISLWDDSLKPCVK VTENFNMWKNNMVEQMHEDIISLWDDSLKPCVK VTENFNMWKNNMVEQMHEDIISLWDDSLKPCVK VTENFNMWKNNMVEQMHEDIISLWDDSLKPCVK VTENFNMWKNNMVEQMHEDIISLWDDSLKPCVK VTENFNMWKNNMVEQMHEDIISLWDDSLKPCVK VTENFNMWKNNNMVEQMHEDIISLWDDSLKPCVK VTENFNMWKNNMVEQMHEDIISLWDDSLKPCVK VTENFNMWKNNMVEQMHEDIISLWDDSLKPCVK VTENFNMWKNNMVEQMHEDIISLWDDSLKPCVK VTENFNMWKNNMVEQMHEDIISLWDDSLKPCVK VTENFNMWKNNMVEQMHEDIISLWDDSLKPCVK VTENFNMWKNNMVEQMHEDIISLWDDSLKPCVK VTENFNMWKNNMMVEQMHEDIISLWDDSLKPCVK VTENFNMWKNNMMVEQMHEDIISLWDDSLKPCVK VTENFNMWKNNMVEQMHEDIISLWDDSLKPCVK VTENFNMWKNNMMVEQMHEDIISLWDDSLKPCVK VTENFNMMKNNMMVEQMHEDIISLWDDSLKPCVK VTENFNMMKNNMMVEQMHEDIISLWDDSLKPCVK			
C1 VTEDFNMWKNDMVEQMHEDIISLWDDSLKPCVK VTENFNMWKNNMVEQMHEDIISLWDDSLKPCVK VTENFNMWKNNNMVEQMHEDIISLWDDSLKPCVK VTENFNMWKNNMVEQMHEDIISLWDDSLKPCVK VTENFNMWKNNMVEQMHEDIISLWDDSLKPCVK VTENFNMWKNNMVEQMHEDIISLWDDSLKPCVK VTENFNMWKNNMVEQMHEDIISLWDDSLKPCVK VTENFNMWKNNMVEQMHEDIISLWDDSLKPCVK VTENFNMWKNNMVEQMHEDIISLWDDSLKPCVK VTENFNMWKNNNMVEQMHEDIISLWDDSLKPCVK VTENFNMWKNNMVEQMHEDIISLWDDSLKPCVK VTENFNMWKNNMVEQMHEDIISLWDDSLKPCVK VTENFNMWKNNMVEQMHEDIISLWDDSLKPCVK VTENFNMWKNNMVEQMHEDIISLWDDSLKPCVK VTENFNMWKNNMVEQMHEDIISLWDDSLKPCVK VTENFNMWKNNMVEQMHEDIISLWDDSLKPCVK VTENFNMWKNNMMVEQMHEDIISLWDDSLKPCVK VTENFNMWKNNMMVEQMHEDIISLWDDSLKPCVK VTENFNMWKNNMVEQMHEDIISLWDDSLKPCVK VTENFNMWKNNMMVEQMHEDIISLWDDSLKPCVK VTENFNMMKNNMMVEQMHEDIISLWDDSLKPCVK VTENFNMMKNNMMVEQMHEDIISLWDDSLKPCVK			
VTEDFNMWKNDMVEQMHEDIISLWDQSLKPCVTENFNMWKNNDMVEQMHEDIISLWDQSLKPCVVTENFNMWKNNMVEQMHEDIISLWDQSLKPCVVTENFNMMVEQMHEDIISLWDQSLKPCVVTENFNMMVEQMHEDIISLWDQSLKPCVVTENFNMMVEQMHEDIISLWDQSLKPCVVTENFNMMVEQMHEDIISLAMPQSLKPCVVTENFNMMVEQMANGONGNAMPGONGONGNAMPGONGONGNAMPGONGONGNAMPGONGONGNAMPGONGONGNAMPGONGONGONGNAMPGONGONGNAMPGONGONGNAMPGONGONGONGONGONGONGONGONGONGONGONGONGONG	ı		
C1- VTEDFNMWKNDDMVEQMHEDIISLWDQSLKPC VTENFNMWKNNMVEQMHEDIISLWDQSLKPC VTENFNMWKNNMVEQMHEDIISLWDQSLKPC VTENFNMWKNNMVEQMHEDIISLWDQSLKPC VTENFNMWKNNMVEQMHEDIISLWDQSLKPC VTENFNMWKNNMVEQMHEDIISLWDQSLKPC VTENFNMWKNNMVEQMHEDIISLWDQSLKPC VTENFNMWKNNMVEQMHEDIISLWDGSLKPC VTENFNMWKNNMVEQMHEDIISLWDQSLKPC VTENFNMWKNNMVEQMHEDIISLWDQSLKPC VTENFNMWKNNMVEQMHEDIISLWDQSLKPC VTENFNMWKNNMVEQMHEDIISLWDQSLKPC VTENFNMWKNNMVEQMHEDIISLWDQSLKPC VTENFNMWKNNMVEQMHEDIISLWDQSLKPC VTENFNMWKNNMVEQMHEDIISLWDQSLKPC VTENFNMWKNNMVEQMHEDIISLWDQSLKPC VTENFNMWKNNMVEQMHEDIISLWDQSLKPC	ı		
C1— VTEDFNMWKNDDMVEQMHEDISLWDQSLKP VTENFNMWKNNMVEQMHEDISLWDQSLKP VTENFNMWKNNMVEQMHEDISLWDQSLKP VTENFNMWKNNMVEQMHEDISLWDQSLKP VTENFNMWKNNMVEQMHEDISLWDQSLKP VTENFNMWKNNMVEQMHEDISLWDQSLKP VTENFNMWKNNMVEQMHEDISLWDGSLKP VTENFNMWKNNMVEQMHEDISLWDGSLKP VTENFNMWKNNMVEQMHEDISLWDQSLKP VTENFNMWKNNMMVEQMHEDISLWDQSLKP VTENFNMWKNNMMVEQMHEDISLWDQSLKP VTENFNMWKNNMMVEQMHEDISLWDQSLKP VTENFNMMKNMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMM			
C1— VTEDFNMWKNDMVEOMHEDIISLWDOSLK VTENFNMWKNNMVEOMHEDIISLWDOSLK VTENFNMWKNNMVEOMHEDIIN	-	1 ⁻	
C1- VTEDFNMWKNDMVEQMHEDIISLWDQSL VTENFNMWKNNMVEQMHEDIISLWDQSL VTENFNMWKNNMVEQMHEDIISLWDQSL VTENFNMWKNNMVEQMHEDIISLWDQSL VTENFNMWKNNMVEQMHEDIISLWDQSL VTENFNMWKNNMVEQMHEDIISLWDQSL VTENFNMWKNNMVEQMHEDIISLWDGSL VTENFNMWKNNMVEQMHEDIISLWDGSL VTENFNMWKNNMVEQMHEDIISLWDGSL VTENFNMWKNNMVEQMHEDIISLWDGSL VTENFNMWKNNMVEQMHEDIISLWDGSL VTENFNMWKNNMVEQMHEDIISLWDGSL VTENFNMWKNNMVEQMHEDIISLWDGSL VTENFNMWKNNMVEQMHEDIISLWDGSL	-		
VTEDFNMWKNDMVEQMHEDIISLWDQS VTENFNMWKNNMWVEQMHEDIISLWDQS VTENFNMWKNNMWVEQMHEDIISLWDQS VTENFNMWKNNMWVEQMHEDIISLWDQS VTENFNMWKNNMVEQMHEDIISLWDQS VTENFNMWKNNMWVEQMHEDIISLWDQS VTENFNMWKNNMWVEQMHEDIISLWDQS VTENFNMWKNNMWVEQMHEDIISLWDQS VTENFNMWKNNMWVEQMHEDIISLWDQS VTENFNMWKNNMWVEQMHEDIISLWDQS VTENFNMWKNNMVEQMHEDIISLWDQS VTENFNMWKNNMVEQMHEDIISLWDQS VTENFNMWKNNMVEQMHEDIISLWDQS VTENFNMWKNNMVEQMHEDIISLWDQS	١	1	
C1 VTEDFNMWKNDMVEQMHEDIISLWDQ VTENFNMWKNNMWVEQMHEDIISLWDQ VTENFNMWKNNNMVEQMHEDIISLWDQ VTENFNMWKNNNMVEQMHEDIISLWDQ VTENFNMWKNNNMVEQMHEDIISLWDQ VTENFNMWKNNNMVEQMHEDIISLWDQ VTENFNMWKNNNMVEQMHEDIISLWDQ VTENFNMWKNNNMVEQMHEDIISLWDQ VTENFNMWKNNNMVEQMHEDIISLWDQ VTENFNMWKNNNMVEQMHEDIISLWDQ VTENFNMWKNNNMVEQMHEDIISLWDQ VTENFNMWKNNMVEQMHEDIISLWDQ VTENFNMWKNNNMVEQMHEDIISLWDQ VTENFNMWKNNMVENNMVEQMHEDIISLWDQ VTENFNMWKNNNMVENNMVEQMHEDIISLWDQ VTENFNMWKNNMVENNMVEQMHEDIISLWDQ VTENFNMWKNNMVENNMVEQMHEDIISLWDQ VTENFNMWKNNMVENNMVEQMHEDIISLWDQ VTENFNMWKNNMVENNMVENNMVENDMHEDIISLWDQ	-		
C1- VTEDFNMWKNDMVEQMHED ISLWD VTENFNMWKNNMVEQMHED ISLWD VTENFNMMWKNNMVEQMHED ISLWD VTENFNMMKNNMVEQMHED ISLWD VTENFNMMWKNNMVEQMHED ISLWD VTENFNMMKNNMVEQMHED ISLWD VTENFNMMKNNMVEQMHED ISLWD VTENFNMMKNNMMVEQMHED ISLWD VTENFNMMMKNNMVEQMHED ISLWD VTENFNMMMKNNMMVEQMHED ISLWD VTENFNMMMKNNMMVEQMHED ISLWD VTENFNMMMMMMMMMMNMMND VTENFNMMMMMMMMMMMMMMMMMMMD VTENFNMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMM	-		!
C1-C1-C1-C1-C1-C1-C1-C1-C1-C1-C1-C1-C1-C	١		
C1-C1-C1-C1-C1-C1-C1-C1-C1-C1-C1-C1-C1-C		222222222222222222222222222222222222222	
C1 VTEDFNMWKNDMVEQMHED VTENFNMWKNNMWVEQMHED VTENFNMWKNNMWVEQMHED VTENFNMWKNNMWVEQMHED VTENFNMWKNNMWVEQMHED VTENFNMWKNNMVEQMHED VTENFNMWKNNMWVEQMHED VTENFNMWKNNMWVEQMHED VTENFNMWKNNMVEQMHED VTENFNMWKNNMWVEQMHED VTENFNMWKNNMWVEQMHED VTENFNMWKNNMVEQMHED VTENFNMWKNNMVEQMHED VTENFNMWKNNMVEQMHED VTENFNMWKNNMVEQMHED VTENFNMWKNNMVEQMHED VTENFNMWKNNMVEQMHED VTENFNMWKNNMVEQMHED VTENFNMWKNNMVEQMHED VTENFNMWKNNNMVEQMHED VTENFNMWKNNMVEQMHED VTENFNMWKNNNMVEQMHED VTENFNMWKNNNMVEQMHED VTENFNMWKNNNMVEQMHED VTENFNMWKNNNMVEQMHED VTENFNMWKNNNMVEQMHED VTENFNMMKNNMVEQMHED VTENFNMMKNNMVEQMHED VTENFNMMKNNMMVEQMHED VTENFNMMKNNMMMVEQMHED VTENFNMMKNNMMMVEQMHED VTENFNMMKNNMMVEQMHED VTENFNMMKNNMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMM	١		
C1—C1—C1—C1—C1—C1—C1—C1—C1—C1—C1—C1—C1—C	ı	S S S S S S S S S S S S S S S S S S S	i
C1—C1—C1—C1—C1—C1—C1—C1—C1—C1—C1—C1—C1—C			
C1—C1—C1—C1—C1—C1—C1—C1—C1—C1—C1—C1—C1—C	١		
C1-C1-C1-C1-C1-C1-C1-C1-C1-C1-C1-C1-C1-C			į
C1-	١	· · · · · · · · · · · · · · · · · · ·	
C1-C1-C1-C1-C1-C1-C1-C1-C1-C1-C1-C1-C1-C		TTTTTT TTTTT	
CTENENEWKKNNWWKNNNWWKNNNWWKNNNWWKNNNWWKNNNNWWKNNNWWKNNNNWWKNNWWKNNNNWWKNNNNWWKNNNNNN	١		
TO STATE OF THE NAME OF THE NA	١		į
TO STATE OF THE NAME OF THE NA	*		
TO STATE OF THE ST	Ĭ		ļ
			ļ
			i
T			
20020000000000000000000000000000000000	١		1
			į
			m
	١		35
>>>>>>>			
>>>>>>>	1		/ 17
>>>>>>			\supseteq
	-	I I	L i
	-	ZZZZZZZZZZZZZZ	

AMENDED SHEET (ARTICLE 19)

11/22

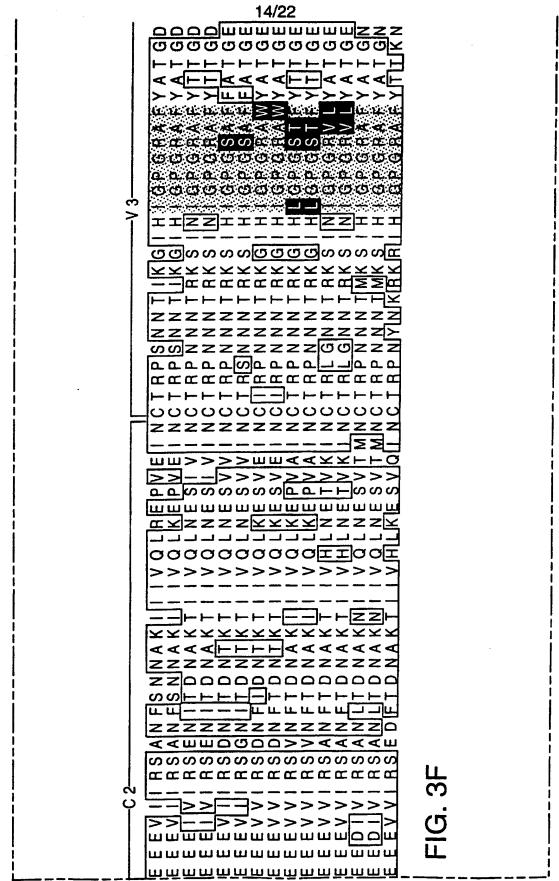
<u>මෙන්නම් මෙන්නම් මෙන්නම් මෙන්න</u>	4.5
TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT	30
╎╶╶┤───────────────────────────────────	\mathfrak{O}
	FIG.
~ ~ ~ ~ ~ ~ ~ ~ ~ ~ ~ ~ ~ ~ ~ ~ ~ ~	\succeq
>	LL.
ZZZZZZZZZZZZZZZZ	
ZZZZŒŒZZZZZZZZ	
スススススススストーススス	
EZZZZZ	
mmcccccccccc	
┆ XXmm∑∑mmm⊓¬¬··X	
ZZMMZZZZ····-	
— — — — — — — — — — — — — — — — — — —	
Q Q Q Q M M Q Q M M Q Q	
M··mm··mm≷≷rr≷≷	
$00 \leftarrow 0022 = 0.00$	
\(\sigma \rightarrow \righ	
ZZZZWWWWWWWUUU	
トトスススストトトースス四四ス	
HHZZHHQQ44HHHZ	
OOZZAAHHHHHHU	
<u> </u>	
- MNZZNONZZH-NONON	
> ZZMMZZXXXQZZZZZ	
44トトトンンののファート	
HHZZZZOOHHHH>>H	
44ZZ <u>HHHHHHH</u> ZZZ	
QQAAAAAZZZZGGH	
0022222F-EE. · -	
NZ H H Q Q K K Q Q H H · · Z	
ZZ> XZ XZ	!
	1
- - · · · · · · · · · · · · · · · · ·	
	j 1
σω	i
σω · · · · · · · · · · · · · · · · · · ·	ļ
ννσσ	
ZZ·····ZZ·····	į
444 444 444 444 444 444 444 444 444 44	; ;
<u> </u>	
7.7.10 63.7.110 63.7.110 63.110	į
MACCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC	
200000000002	

AMENDED SHEET (ARTICLE 19)

	12/22	
	HH4444444444	
	TTTTTTTTTTT	
Ġ.	<u>⊢</u>	
Ϋ́	<u> </u>	
	XXXXXXXXXXXX	
ľ		
[44444444444	
	3000000000000	
	 	
	>>>>>>>>>>>>	
	ZZZZZZZZZZZZZ	
Γ		
	S S S S S S S S S S S S S S S S S S S	
		İ
	>> <u>>>></u> >>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>	
	∾∾ <u>SS⊢⊢</u> ∾∾∾∾∾∾∾∾∾	
Ī		
	ZZ · ·ZZ · ·OOOOZZZ	
]		! !
	· · · · · · · · · · · · · · · · · · ·	
	a a O O a a a a a a a a a	
Ì	>>>>>>>>>	1
	>>>>>	
		į
1] [
	N N X X K K K X X X X X X X X X X X X X	Į
	<u> </u>	į
ļ		i !
1		
7.5	>>>==>>>> M	į
1		! !
	*******	į
1	>>>QXZ4>>>rrZ	!
	スペスススススススススは日内内	ŀ

13/22

44444444**4**444 **აააააააააააააა ტიტიტიტიტიტიტიტი** gaaaaaaaaaaaaa - }- }- }- }- }o>>>>>>>>>> >>>>>>>>>> - -- -- -- -- -- -- -- --**Ა Ა Ა|Ლ|Ა Ა Ა Ა Ა Ა Ა Ა Ა Ა Ა** 0000000<u>00</u>00<u>0</u>0 - トニススススートトト **ZZZWWWWZZZ** <u>CXXX</u>ZZZZXXZZXX 、ススススススススススス *4444444444444* **444444444444** C6.1 C6.5 C8.3 C8.6 C15.2 C7.2 C7.10 C11.5 C10.5 C17.3 MNGNE



15/22

	$\overline{\alpha\alpha\alpha\alpha\alpha\alpha\alpha\alpha\alpha\alpha\alpha\alpha\alpha}$
	IIXXCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC
1	
i	>>44>>00-44>>>
ļ	
Ì	ZZZZMMZZZZZZZZZZZZZZZZZZZZZZZZZZZZZZZZ
	₽₽ · QQ · QQUU · ·
1	· · · · · · · · · · · · · · · · · · ·
İ	· · · · · gg · · · · · · · ·
	··************************************
1	<u> </u>
	aakkaakkaaaaaa
1	**************************************
	\ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \
l l	> ⊼ETTERZ ZZ ⊼⊼ QŒŢŢŢ
	<u>>><<>>><<</u> >>
	<u> </u>
	AXECQQEEMMAXE
į	
,	
1	N N N N N N N N N N N N N N N N N N N
1	
	**AGGGGXXXX
	AAFFFFFFFFFAAAFFA
	\bigcirc
į	ZZZZZZZYYZZZZ
i	0000000 <u>00</u> 0000 <u>0</u>
1	QQQQQQQQQQ
i	
	5
1	3323 3323 3323 3323 3323 3323 3323 332
1	~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
1	ய
	7.10 11.5 11.5 11.5 11.7 11.7 11.7 11.3 11.3 11.3 11.3 11.3
	88882222222288888
i	

16/22	
HHZZZZZHHZZHHZ	
NOUDONNOOUN	
ZZ · · · · - ⊢ - Q Q · · · · · · · Q	
A MMZOSSOSSOSSISTELE	
> H-QQQQAAXXZZZ	
· · · · · · · · · · · · · · · · · · ·	
· ·	
· ·ZZZZZZ · · · · · · · Z	
··00ZZ\$\$ · · · · · ·	
· · · · · · · · · · · · · · · · · · ·	
· · • • • • • • • • • • • • • • • • • •	
OSSZZZCC<	
ZZZZZZZZZZZZZZZ	
333333333333333333333333333333333333333	
\circ	
ZZZZZZZZZZZZZZ	
XXXXAAQQQQQQQQ	
FFFFFFFFFFF	
ZZZØZZOZZZZZZZZ	
>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>	
00000000000000	!
<u> </u>	
000000000000000	
ZZZZZZZZZZZZZZZ	į
	i i
N	
エエエエエエエエエ	i
>>>>\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\	!
	I
ө пппппппппппппппп пп	ह
γασασασασασασα	
	U
000000000000000	
	

17/22 OOOOMMZZHHOOSSOHHHHZZZZ>>OOZZH ZZZZYYWWZZOOZZO 31 SOSSZEEZZZZ დიდიდიდიდიდი α **იიიიიიიიიიიიიი** ZZZZZZZZZZZZZZ ∞ **∾∾∾∾∾∾∾** 00000000<u>00</u>00000 XX**月月月月SSSSHHHHKKKKI 44444444444** ¥₹₹₹₹₹₹₹ **444444444444** 00000000000000 aaaaaaaaaaaaa 427 427 414 C6.5 C6.5 C8.3 C8.6 C15.2 C7.2 C7.10 C11.5 C10.5 C17.1 C10.5 MNGNE

	18/22					
GEEREE E						
	Z XXXXX					
REEREE E						
aaaaaaa						
>>>>>>>>>>						
A A A A A A A A A A A A A A A A A A A						
######################################						
XXXXXX	D					
AAAAAAAA						
XXQQXXXDX						
	$\vdash\vdash\vdash\vdash\vdash\vdash$					
	F					
>>>>>						
<u> </u>	P.					
		38	30	3F	픙	7
				1 .		i. 3J
·	>	<u> 일</u>	<u> </u>	B B	FIG.	FIG
		-				
>>>>>>	 -	¥	ဗ္က	38	36	ਲ
>>>>>>		<u> </u>	FIG.	<u>ত</u>	FIG.	FIG.
スススススススス	スススススス	☲	正	프	田	ᄑ
> > > > > > > > > > > > > > > > >				·		3
						E E
>>>>>>>	•					正
S S S S S S S S S S S S S S S S S S S						
ZZZZZZZZ						
THE THE THE THE						
2222222	ZZZZZ					•
00000000	00000					
>>						
	· C C · · · ·		_)		
<u>,</u>	·ZZ···		(1	5		
	· · · ·					
	· · · ·		٢	5		
· · Œ Œ ⊢ · · · · · · · · · · · · · · ·				_		
 4422 mm2	ZOOXXZ		Ц	-		
OOOOWZZZ	ZZZWWH					

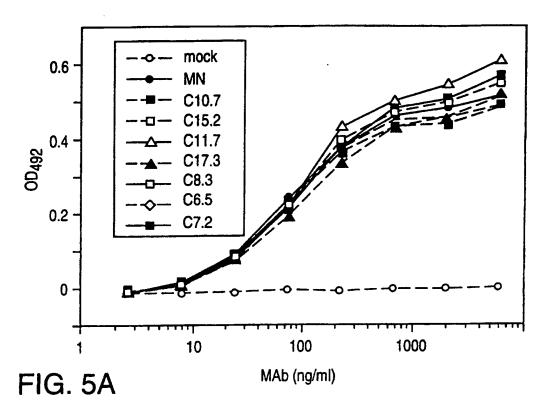
19/22

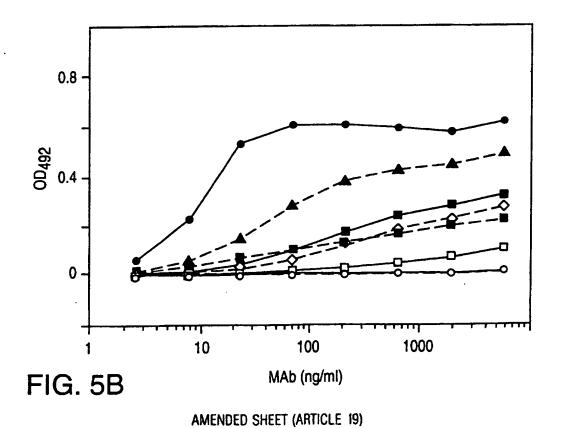
a b				
C10 a b				
0 c				
C17 a b				رن
٤				CT.
Z				
C7				
0 g			:	
82 0 e			 .·	
	<u>kDa</u> 200 –	- (5)	- 99	

AMENDED SHEET (ARTICLE 19)

WO 98/01564 PCT/US97/09690

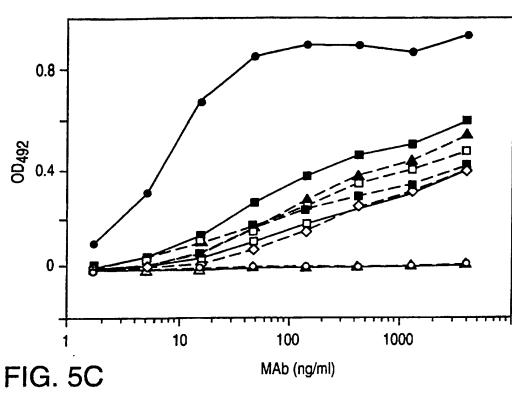


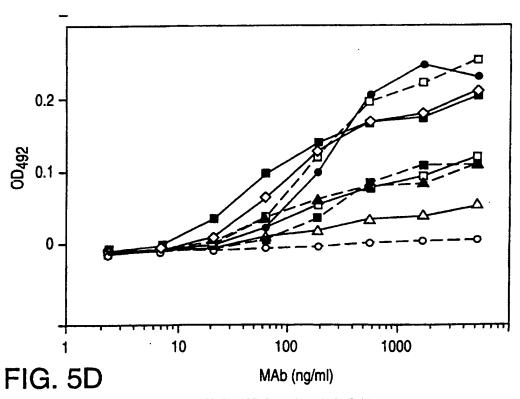




PCT/US97/09690 WO 98/01564







AMENDED SHEET (ARTICLE 19)

WO 98/01564 PCT/US97/09690

22/22

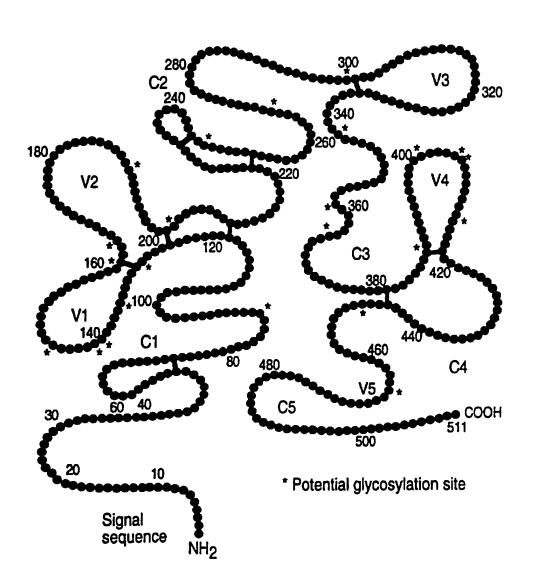


FIG. 6

INTERNATIONAL SEARCH REPORT

enterr nat Application No PCT/US 97/09690

A CLASSIFICATION OF SUBJECT MATTER IPC 6 C12N15/49 C07 C07K14/16 A61K39/21 According to International Patent Classification (IPC) or to both national classification and IPC 8. FIELDS SEARCHED Minimum documentation searched (classification system followed by classification symbols) C12N C07K A61K IPC 6 Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched Electronic data base consulted during the international search (name of data base and, where practical, search terms used) C. DOCUMENTS CONSIDERED TO BE RELEVANT Relevant to claim No. Citation of document, with indication, where appropriate, of the relevant passages Category * 1-18, WO 94 28929 A (GENETECH, INC.) 22 December X 20-25 1994 see page 56 SEQ. ID. NO. 25. see page 50, line 14 - line 31 1,5,6 P.W. BERMAN ET AL.: "Genetic and Χ immunologic characterization of viruses infecting MN-rgp120 vaccinated volunteers" ONE WORLD, ONE HOPE: XI INTERNATIONAL CONFERENCE ON AIDS, vol. 10, no. supplement 3, 7 - 12 July 1996, VANCOUVER, CANADA, page 10 XP002045307 See "Methods" in Abstract Mo.A.285 -/--Patent family members are fisted in annex. Further documents are listed in the continuation of box C. * Special categories of cited documents : "I" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the *A* document defining the general state of the art which is not considered to be of particular relevance "X" document of perticular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone "E" earlier document but published on or after the international filing date "L" document which may throw doubts on priority claim(s) or which is claid to establish the publication date of another citation or other special reason (as appealed) "Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled *O* document referring to an oral disclosure, use, exhibition or in the art. *P" document published prior to the international filing data but later than the priority date claimed "&" document member of the same patent family Date of mailing of the international search report Date of the actual completion of the international search 2 6 -11- 1997 30 October 1997 Authorized officer Name and mailing address of the ISA European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl Cupido, M Fax: (+31-70) 340-3016

1

INTERNATIONAL SEARCH REPORT

Inter and Application No PCT/US 97/09690

0.10==1===	Man LOCCHIENTO CONSIDERAN	PCT/US 97/09690
C.(Continua Category *	ation) DOCUMENTS CONSIDERED TO BE RELEVANT Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
		775572.1. D 06111110.
A	M.J. MCELRATH ET AL.: "Human immunodeficiency virus type 1 infection despite prior immunization with a recombinant envelope vaccine regimen" PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES OF USA, vol. 93, no. 9, 30 April 1996, WASHINGTON US,	19
	pages 3972-3977, XP002045308 see page 3976, last paragraph; figure 1	
Τ	P.W. BERMAN ET AL.: "Genetic and immunologic characterization of viruses infecting MN-rgp120-vaccinated volunteers"	1-25
	THE JOURNAL OF INFECTIOUS DISEASES, vol. 176, no. 2, August 1997, pages 384-397, XP002045309 see the whole document	
-		
į.		
ĺ		

1

INTERNATIONAL SEARCH REPORT

Inter mai Application No

	stormation or natural familie members		Inter Mai Application No		
information on patent family members		PCT/US 97/09690			
Patent document cited in search report	Publication date	Patent family member(s)		Publication date	
WO 9428929 A	22-12-94	AU 7047894 EP 070865	4 A 9 A	03-01-95 01-05-96	
	:				
	¥				
	,				